



PN WO200056762-A2.  
 XX 28-SEP-2000.  
 PD 22-MAR-2000; 2000WO-US07781.  
 XX 22-MAR-1999; 99US-0273623.  
 PR 22-MAR-1999; 99US-0273623.  
 XX (NOVO ) NOVO NORDISK BIOTECH INC.  
 PA (NOVO ) NOVO NORDISK AS.  
 XX Berka RM, Rey MM, Shuster JR, Kaupinen S, Clausen IG, Olsen PB;  
 PI WPI; 2000-594572/56.  
 DR  
 XX  
 PT Monitoring differential expression of genes in filamentous fungal cells  
 PT uses fluorescence-labeled nucleic acids isolated from the cells and a  
 PT substrate of expressed sequence tags -  
 XX  
 XX  
 XX Claim 86; Page 550; 3161pp; English.  
 CC The present invention describes a method for monitoring differential  
 CC expression of genes in a first filamentous fungal (FF) cell relative to  
 CC expression of the same genes in one or more second filamentous fungal  
 CC cells. The method uses fluorescence-labeled nucleic acids isolated from  
 CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs  
 CC are used in the methods for monitoring differential expression of genes  
 CC in a first filamentous fungal (FF) cell relative to expression of the  
 CC same genes in one or more second filamentous fungal cells. Monitoring  
 CC the global expression of genes from FF cells allows the production  
 CC potential of the microorganisms to be improved. New genes may be  
 CC discovered, possible functions of unknown open reading frames can be  
 CC identified and gene copy number variation and stability can be  
 CC monitored. The expression of genes can be used to study how FF cells  
 CC adapt to changes in culture conditions, environmental stresses, spore  
 CC morphogenesis, recombination, metabolic or catabolic pathway  
 CC engineering. Using ESTs provides several advantages over genomic or  
 CC random DNA clones including elimination of redundancy as one spot on an  
 CC array equals one gene or open reading frame, and organization of the  
 CC microarrays based on function of the gene products to facilitate  
 CC analysis of the results. AAF07478 to AAF11247 represents ESTs from  
 CC *Aspergillus fumigatus*; AAF11248 to AAF11853 represents ESTs from *Aspergillus*  
 CC *niger*; AAF11854 to AAF14878 represents ESTs from *Aspergillus oryzae*; and  
 CC AAF14879 to AAF15337 represents ESTs from *Trichoderma reesei*, which are  
 CC all specifically claimed in the present invention.  
 XX  
 XX  
 SQ Sequence 631 BP; 125 A; 223 C; 141 G; 134 T; 8 other;  
 XX  
 XX Alignment Scores:  
 XX Pred. No.: 3.6e-05 Length: 631  
 XX Score: 81.00 Matches: 15  
 XX Percent Similarity: 75.00% Conservative: 6  
 XX Best Local Similarity: 53.57% Mismatches: 7  
 XX Query Match: 55.86% Indels: 0  
 XX DB: 21 Gaps: 0  
 XX  
 US-09-987-190-2 (1-30) x AAF07909 (1-631)  
 QY 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSeryI 20  
 DB 123 AAGCGCAGCTCTCCGAGCTCTTACNACTACGCGCTCTTGAGCCCTACATCTCTGGC 182  
 QY 21 GlnIleAsnGluIle\*\*TyrThr 28  
 DB 183 CAGATCATGAGAGCTCCACCACTCC 206  
 XX  
 XX RESULT 2  
 XX AAT85877  
 XX ID AAT85877 standard; cDNA to mRNA; 812 BP.  
 XX AC AAT85877;  
 XX DT 23-FEB-1998 (first entry)

XX  
 XX Malassezia fungus MF-4 antigenic protein encoding cDNA.  
 XX  
 XX Malassezia; fungus; antigenic; human; IGE; immunoglobulin E;  
 KW antibody; allergy; antigen; ds.  
 XX  
 XX Malassezia sp.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 XX FT 2..676  
 XX FT CDS /tag= a  
 XX /product= MF-4\_antigenic\_protein  
 XX  
 XX WO9721817-A1.  
 XX  
 XX 19-JUN-1997.  
 XX  
 XX 10-DEC-1996; 96WO-JP03602.  
 XX  
 XX 05-SEP-1996; 96JP-0257613.  
 XX 12-DEC-1995; 95JP-0346627.  
 XX 05-SEP-1996; 96JP-0257612.  
 XX  
 XX (TAKI ) TAKARA SHUZO CO LTD.  
 XX  
 XX Akiyama K, Kato I, Kuroda M, Okado T, Onishi Y;  
 PI Takesako K, Yagihara T, Yamaguchi H, Yasueda H;  
 XX WPI; 1997-332788/30.  
 DR P-PSDB; AAW2977L.  
 DR  
 XX  
 XX Antigenic proteins from the fungus Malassezia - bind to IGE  
 PT antibodies present in patients with Malassezia allergies, useful for  
 PT diagnosis, treatment and prevention of such conditions  
 XX  
 XX  
 XX Claim 56; Page 77-78; 162pp; Japanese.  
 XX  
 XX The present sequence encodes a specifically claimed antigenic  
 CC protein isolated from the fungus Malassezia. The antigenic protein  
 CC can bind to IGE antibodies present in patients with allergic  
 CC conditions. Antigenic proteins, peptides and nucleic acids from the  
 CC fungus Malassezia can be used in the diagnosis, treatment and  
 CC prevention of allergic conditions due to Malassezia organisms (such  
 CC as M. furfur, M. sympodialis and M. pachydermatis).  
 XX  
 XX  
 SQ Sequence 812 BP; 235 A; 245 C; 218 G; 114 T; 0 other;  
 XX  
 XX Alignment Scores:  
 XX Pred. No.: 0.000115 Length: 812  
 XX Score: 79.00 Matches: 15  
 XX Percent Similarity: 70.37% Conservative: 4  
 XX Best Local Similarity: 55.56% Mismatches: 8  
 XX Query Match: 54.48% Indels: 0  
 XX DB: 18 Gaps: 0  
 XX  
 US-09-987-190-2 (1-30) x AAT85877 (1-812)  
 QY 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSeryI 20  
 DB 74 AAGTACAGCGCTCCGCGCTGCGCTACGACTACGCGCGCTCGAGCCGCGACTCTCGGCG 133  
 QY 21 GlnIleAsnGluIle\*\*Tyr 27  
 DB 134 GAGATCATGAGAGAGCACTAC 154  
 XX  
 XX RESULT 3  
 XX AAH68192  
 XX ID AAH68192 standard; DNA; 600 BP.  
 XX AC AAH68192;  
 XX DT 26-SEP-2001 (first entry)



DE C glutamicum coding sequence fragment SEQ ID NO: 3227.  
 XX  
 XX Corynebacterium; amino acid synthesis; vitamin; saccharide;  
 KM organic acid synthesis; ds.  
 XX  
 OS Corynebacterium glutamicum.  
 XX  
 PN EP1108790-A2.  
 XX  
 PD 20-JUN-2001.  
 XX  
 PF 18-DEC-2000; 2000EP-0127688.  
 XX  
 PR 16-DEC-1999; 99JP-0377484.  
 PR 07-APR-2000; 2000JP-0159162.  
 PR 03-AUG-2000; 2000JP-0280988.  
 XX  
 PA (KYOW ) KYOWA HAKKO KOGYO KK.  
 XX  
 XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;  
 XX Tateishi N, Senoh A, Ikeda M, Ozaki A;  
 XX WPI; 2001-376931/40.  
 DR P-PSDB; AAG92973.  
 XX  
 XX Novel polynucleotides derived from Corynebacterium bacteria, for identifying  
 PT mutation point of a gene, measuring expression of a gene, analysing  
 PT expression profile or pattern of a gene and identifying homologous gene  
 PT  
 XX  
 PS Claim 8; SEQ ID NO: 3227; 246bp + Sequence Listing; English.  
 XX  
 XX The present invention provides a number of nucleotide and protein  
 CC sequences from the Corynebacterium bacterium Corynebacterium glutamicum. These  
 CC are useful for identifying the mutation point of a gene derived from a  
 CC mutant of corynebacterium bacterium, measuring expression amount and  
 CC analysing the expression profile or expression pattern of a gene derived  
 CC from Corynebacterium bacterium, and identifying a homologue of a gene derived  
 CC from corynebacterium bacterium. Corynebacterium bacteria are useful for producing  
 CC amino acids, nucleic acids, vitamins, saccharides and organic acids,  
 CC particularly L-lysine. The present sequence is a nucleic acid described  
 CC in the exemplification of the invention.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from the  
 CC European Patent Office.  
 CC  
 XX  
 SO Sequence 600 BP; 143 A; 197 C; 146 G; 114 T; 0 other;  
 XX  
 XX Alignment Scores:  
 XX Pred. No.: 0.00018 Length: 600  
 XX Score: 77.00 Matches: 13  
 XX Percent Similarity: 74.07% Conservative: 7  
 XX Best Local Similarity: 48.15% Mismatches: 7  
 XX Query Match: 53.10% Indels: 0  
 XX DB: 22 Gaps: 0  
 XX  
 US-09-987-190-2 (1-30) x AAH68192 (1-600)  
 QY 2 TySerLeuProGluLeuAspTYrGluPheserAlaThrGluProTYrIleSerGlyIn 21  
 DB 10 TACGAACTCCCGAAGACTGACTACGACCTCTCGAGCCACATCGCGCTGAA 69  
 QY 22 IleAngGluIle\*\*\*TYrThr 28  
 DB 70 ATCATGGAGCTTACCACTCC 90  
 RESULT 4  
 ID AAS00520 standard; DNA; 1143 BP.  
 AC AAS00520;  
 XX  
 DT 09-MAY-2001 (first entry)

XX  
 DE C. melassecola superoxide dismutase DNA.  
 XX  
 XX Superoxide dismutase; sod; corynebacterium microorganism; metabolite; ds;  
 KM vitamin; D-panthothenic acid; L-lysine; amplification; animal nutrition.  
 XX  
 OS Corynebacterium melassecola.  
 XX  
 PN EP1077261-A2.  
 XX  
 PD 21-FEB-2001.  
 XX  
 PF 02-AUG-2000; 2000EP-0116669.  
 XX  
 PR 13-AUG-1999; 99US-0373731.  
 XX  
 PA (DEGS ) DEGUSSA-HUELS AG.  
 XX (CNRS ) CNRS CENT NAT RECH SCI.  
 XX  
 XX Merkamm M, Guyonvarch A, Marx A;  
 XX WPI; 2001-212718/22.  
 DR P-PSDB; AAU00514.  
 XX  
 XX New DNA encoding superoxide dismutase of Corynebacterium, useful for  
 PT producing transformants with increased production of metabolites,  
 PT particularly lysine -  
 PT  
 XX  
 PS Claim 1; Page 13-14; 19pp: English.  
 XX  
 XX The sequence represents Corynebacterium melassecola superoxide dismutase  
 CC (sod) DNA. Corynebacterium microorganisms may be transformed with a sod DNA  
 CC sequence and the sod gene can be amplified. Sod is then often  
 CC overexpressed in corynebacterium bacteria. Corynebacterium that overexpress sod  
 CC are used for production of metabolites, particularly nucleotides,  
 CC vitamins and amino acids, especially D-panthothenic acid or, specifically,  
 CC L-lysine. These metabolites are useful in human or animal nutrition and  
 CC as pharmaceuticals. Overexpression of sod in corynebacterium increases the  
 CC yield of particular metabolites.  
 CC  
 XX  
 SO Sequence 1143 BP; 273 A; 323 C; 288 G; 259 T; 0 other;  
 XX  
 XX Alignment Scores:  
 XX Pred. No.: 0.000409 Length: 1143  
 XX Score: 77.00 Matches: 13  
 XX Percent Similarity: 74.07% Conservative: 7  
 XX Best Local Similarity: 48.15% Mismatches: 7  
 XX Query Match: 53.10% Indels: 0  
 XX DB: 22 Gaps: 0  
 XX  
 US-09-987-190-2 (1-30) x AAS00520 (1-1143)  
 QY 2 TySerLeuProGluLeuAspTYrGluPheserAlaThrGluProTYrIleSerGlyIn 21  
 DB 347 TACGAACTCCCGAAGACTGACTACGACCTCTCGAGCCACATCGCGCTGAA 406  
 QY 22 IleAngGluIle\*\*\*TYrThr 28  
 DB 407 ATCATGGAGCTTACCACTCC 427  
 RESULT 5  
 ID AAH68534 standard; DNA; 309400 BP.  
 AC AAH68534;  
 XX  
 DT 26-SEP-2001 (first entry)

DE C glutamicum coding sequence fragment SEQ ID NO: 7069.  
 XX Corynebacterium; amino acid synthesis; vitamin; saccharide;  
 KW organic acid synthesis; ds.  
 XX Corynebacterium glutamicum.  
 OS  
 PN BPI108790-A2.  
 XX  
 PD 20-JUN-2001.  
 XX  
 PF 18-DEC-2000; 2000EP-0127688.  
 XX  
 PR 16-DEC-1999; 99JP-0377484.  
 PR 07-APR-2000; 2000JP-0159152.  
 PR 03-AUG-2000; 2000JP-0280988.  
 XX  
 PA (RYOW ) KYOWA HAKKO KOGYO KK.  
 XX  
 PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;  
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;  
 DR WPI; 2001-376931/40.  
 XX  
 PT Novel polynucleotides derived from Corynebacterium bacteria, for identifying  
 PT mutation point of a gene, measuring expression of a gene, analysing  
 PT expression profile or pattern of a gene and identifying homologous gene  
 PT  
 PS Disclosure; SEQ ID NO: 7069; 246bp + Sequence Listing; English.  
 XX  
 CC The present invention provides a number of nucleotide and protein  
 CC sequences from the Corynebacterium bacterium Corynebacterium glutamicum. These  
 CC are useful for identifying the mutation point of a gene derived from a  
 CC mutant of Corynebacterium bacterium, measuring expression amount and  
 CC analysing the expression profile or expression pattern of a gene derived  
 CC from Corynebacterium bacterium, and identifying a homologue of a gene derived  
 CC from Corynebacterium bacterium. Corynebacterium bacteria are useful for producing  
 CC amino acids, nucleic acids, vitamins, saccharides and organic acids,  
 CC particularly L-lysine. The present sequence is a nucleic acid described  
 CC in the exemplification of the invention.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from the  
 CC European Patent Office.  
 CC  
 XX  
 SQ Sequence 309400 BP; 70133 A; 86477 C; 83115 G; 69675 T; 0 other;  
 XX  
 Alignment Scores:  
 e: No: 0.528 Length: 309400  
 Percent Similarity: 77.00% Matches: 13  
 Best Local Similarity: 48.15% Conservative: 7  
 Query Match: 53.10% Mismatches: 7  
 DB: 22 Gaps: 0  
 US-09-987-190-2 (1-30) x AAH68534 (1-309400)  
 QY 2 TyrserrleupProGluLeuAspTyrGluPheserAlarhGluProTyrIleSerGlyIn 21  
 DB 126401 TACCAACTCCGAGAACTGACACTACGACATGACGACATGCGCGTGA 126460  
 QY 22 IleAsngluile\*\*TyrThr 28  
 DB 126461 ATCATGAGCTTCACCACTCC 126481  
 RESULT 6  
 AAQ94270  
 ID AAQ94270 standard; cDNA; 97 BP.  
 XX  
 AC AAQ94270;  
 XX  
 DT 24-MAY-1996 (first entry)  
 XX

DE Human Mn-superoxide dismutase (lys29) N-terminal coding sequence.  
 XX  
 KW hMn-SOD; manganese; superoxide dismutase; terramer; inflammation;  
 KW autoimmune disease; rheumatoid arthritis; neoplasm; emphysema;  
 KW osteoarthritis; wound healing; ds.  
 XX  
 OS Synthetic.  
 XX  
 FH Key location/Qualifiers  
 FT mat\_peptide 4..97  
 FT /tag= a  
 FT /note= "only encodes partial mature protein  
 sequence, i.e. N-terminal amino acids 1-31"  
 XX  
 PN EP676472-A1.  
 XX  
 PD 11-OCT-1995.  
 XX  
 PF 10-MAR-1988; 88EP-0107460.  
 XX  
 PR 24-DEC-1987; 87DE-3744038.  
 PR 14-MAR-1987; 87DE-3708306.  
 PR 26-MAY-1987; 87DE-3717695.  
 PR 10-JUL-1987; 87DE-3722884.  
 XX  
 PA (BOEH ) BOEHRINGER INGELHEIM INT GMBH.  
 XX  
 PI Hauptmann R, Heckl K, Krystek E, Maurer-Fogy I;  
 PI Ostermann E, Spevak W, Stralowa C, Wiche-Castanon M;  
 PI Zoephel A;  
 XX  
 DR WPI; 1995-346092/45.  
 XX  
 PT Genes encoding recombinant human manganese superoxide dismutase -  
 PT for treatment, prevention and diagnosis of inflammatory diseases  
 PT  
 PS Claim 8; Page 5; 54pp; German.  
 XX  
 CC A human placental cDNA library was screened with a probe derived  
 CC from published sequences of human liver MnSOD. Two sequences were  
 CC isolated which encoded the majority (either amino acids 22-198 or  
 CC 26-198) of hMnSOD. To complete the 5'-ends of the cDNA sequences,  
 CC a double-stranded fragment (AAQ94270) coding for an initial Met  
 CC residue followed by amino acids 1-31 was synthesised using yeast  
 CC preferred codons. Recombinant, full-length hMn-SOD is useful for  
 CC treating, preventing or diagnosing inflammatory, degenerative,  
 CC neoplastic and rheumatic disorders; in wound healing and in  
 CC autoimmune diseases.  
 CC  
 XX  
 SQ Sequence 97 BP; 30 A; 30 C; 15 G; 22 T; 0 other;  
 XX  
 Alignment Scores:  
 e: No: 6.12e-05 Length: 97  
 Score: 74.00 Matches: 12  
 Percent Similarity: 78.57% Conservative: 10  
 Best Local Similarity: 42.86% Mismatches: 6  
 Query Match: 51.03% Indels: 0  
 DB: 16 Gaps: 0  
 US-09-987-190-2 (1-30) x AAQ94270 (1-97)  
 QY 1 LysTyrSerleupProGluLeuAspTyrGluPheserAlarhGluProTyrIleSerGly 20  
 DB 4 AAGCACTCTTGGCCAGACTTGCCATACGACTACGCGTCTTAAGCAACACATCAATGCT 63  
 QY 21 GlnIleAsngluile\*\*TyrThr 28  
 DB 64 CAATCATGCAATGACCACTCT 87  
 RESULT 7  
 AAC02178  
 ID AAC02178 standard; cDNA; 414 BP.  
 XX

AAC02178;  
 06-OCT-2000 (first entry)  
 Human secreted protein 5' EST, SEQ ID NO: 2176.  
 Human: 5' EST; expressed sequence tag; secreted protein; cDNA isolation, gene therapy; chromosome mapping; ss.  
 Homo sapiens.  
 EPI033401-A2.  
 06-SEP-2000.  
 21-FEB-2000; 2000EP-0200610.  
 26-FEB-1999; 99US-0122487.  
 (GEST ) GENSET.  
 Dunas Milne Edwards J, Duclert A, Giordano J;  
 WPI: 2000-500381/45.  
 P-PSDB: AAG02172.  
 New nucleic acid that is a 5' expressed sequence tag (5' EST) for obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for diagnostic, forensic, gene therapy and chromosome mapping procedures -  
 Claim 1; SEQ ID 2176; 71pp + CD-ROM, English.  
 The present sequence is one of a large number of 5' ESTs derived from mRNAs encoding secreted proteins. An ORF has been identified within the sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs derived from 30 different tissues. EST sequences usually correspond mainly to the 3' untranslated region (UTR) of the mRNA because they are often obtained from oligo-dT primed cDNA libraries. Such ESTs are not well suited for isolating cDNA sequences derived from the 5' ends of mRNAs and even in those cases where longer cDNA sequences have been obtained, the full 5' UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and chromosome mapping procedures. They are used to obtain upstream regulatory sequences and to design expression and secretion vectors.  
 Sequence 414 BP, 89 A, 124 C, 126 G, 73 T, 2 other;  
 Alignment Scores:  
 Pred. NO.: 0.000391 Length: 414  
 Score: 74.00 Matches: 12  
 Percent Similarity: 78.57% Conservative: 10  
 Best Local Similarity: 42.86% Mismatch: 6  
 Query Match: 51.03% Indels: 0  
 Gaps: 21  
 US-09-987-190-2 (1-30) x AAC02178 (1-414)  
 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSercly 20  
 182 AAGCACACGCTCCCGACCTGCGCTACGACGAGCGCCCTGGAACCTCATCAGCGG 241  
 21 GlnIleAsnGluIle\*\*TyrThr 28  
 242 CAGATCATGCAGCTGCACACGAGC 265  
 RESULT 8  
 AAZ08755  
 AAZ08755 standard; cDNA, 552 BP.  
 AAZ08755;

```

DT 03-NOV-1999 (first entry)
XX
DE Human manganese superoxide dismutase exon 3-deleted isoform cDNA.
XX
KW Human; manganese superoxide dismutase; MnSOD; exon 3 deleted isoform;
KW MnSOD E3(-); diagnosis; viral infection; HIV; oncogenesis; tumour;
KW UV-induced damage; post ischaemia reperfusion damage; anti-inflammatory;
KW cytotoxic; cardiotoxic; cancer; ss.
XX
OS Homo sapiens.
OS Synthetic.
XX
FT Key Location/Qualifiers
FT CDS 1..552
FT /tag= a
FT /product= "MnSOD E3(-)"
FT /note= "manganese superoxide dismutase exon 3-deleted
FT isoform"
XX
PN M09943697-A1.
XX
PD 02-SEP-1999.
XX
PF 25-FEB-1999; 99MO-US04129.
XX
PR 25-FEB-1998; 98US-0075948.
XX
PA (UYCO-) UNIV & COMMUNITY COLLEGE.
XX
PI Anziano PQ;
XX
DR WPI; 1999-527592/44.
XX
DR P-PSDB; AAY29656.
XX
PT A new isoform useful for diagnosing oxidative stress, and treating
PT viral infections
XX
PS Claim 2; Page 12; 64pp; English.
XX
XX
CC The present sequence encodes an isoform of manganese superoxide
CC dismutase (MnSOD), which is a splice-variant lacking exon 3 of the full
CC length MnSOD (i.e. MnSOD E3(-)). MnSOD E3(-) may be useful in the
CC treatment of a wide variety of disorders including viral infections,
CC particularly HIV, and may be used for the prevention of oncogenesis,
CC tumour promotion and invasiveness, and UV-induced damage, for protection
CC of cardiac tissue against post ischaemia reperfusion damage, as an
CC anti-inflammatory agent, to reduce the cytotoxic and cardiotoxic effects
CC of anti-cancer drugs, and to improve the longevity of living cells.
XX
SQ Sequence 552 BP; 145 A; 135 C; 153 G; 119 T; 0 other;
XX
XX
Alignment Scores:
Pred. No.: 0.000565 Length: 552
Score: 74.00 Matches: 12
Percent Similarity: 78.57% Conservative: 10
Best Local Similarity: 42.86% Mismatches: 6
Query Match: 51.03% Indels: 0
DB: 20 Gaps: 0
US-09-987-190-2 (1-30) x AAZ08755 (1-552)
QY 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSergly 20
Db 73 AAGCAGACGCGCTCCCGACGCTGCCCTTACGATCAGCGCGCCCTGGACCTCATCAGCGG 132
QY 21 GlnIleGluIle***TyrThr 28
Db 133 CAGATCATGCGAGCTGCACACACAGC 156
RESULT 9
ID AAQ73585
XX AAQ73585 standard; DNA; 561 BP.
XX

```

```

AC      AAQ73585;
XX
XX      25-JUN-1995 (first entry)
XX
XX      Human manganese superoxide dismutase variant gene.
DE
XX      MnSOD; tetramer; pulmonary oxygen toxicity; mitochondrial; mutant;
XX      88.
XX
XX      Homo sapiens.
OS
XX      MO9421283-A.
XX
XX      29-SEP-1994.
PD
XX
XX      24-MAR-1994; 94WO-US03185.
XX
XX      24-MAR-1993; 93US-0036604.
XX
XX      (BOEH ) BOEHRINGER INGELHEIM PHARM INC.
XX
XX      Wegner C, Wolyniec W;
XX
XX      WPI; 1994-316659/39.
XX
XX      Improved inhibition of pulmonary oxygen toxicity - by
XX      prophylactic, topical admin. of human mitochondrial manganese
XX      superoxide dismutase
XX
XX      Claim 7; Page 26; 41pp; English.
XX
XX      The sequence is that of a mutant human manganese superoxide
XX      dismutase tetramer. Mutations of the native protein provide
XX      a protein which, when applied topically, can inhibit pulmonary
XX      oxygen toxicity.
XX      See also AAQ73582-4.
XX
XX      Sequence 561 BP, 159 A, 127 C, 139 G, 136 T, 0 other;
SQ
XX
XX      Alignment Scores:
XX      Pred. No.: 0.000577 Length: 561
XX      Score: 74.00 Matches: 12
XX      Percent Similarity: 78.57% Conservative: 10
XX      Best Local Similarity: 42.85% Mismatch: 6
XX      Query Match: 51.03% Indels: 0
XX      DB: 15 Gaps: 0
XX
XX      US-09-987-190-2 (1-30) x AAQ73585 (1-561)
XX
XX      1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSergly 200
XX      4 AAGCACTCTTTGCCAGACTTGCATTCGACTACGCGTCTCTGAGAACACACATCATGCT 630
XX
XX      21 GlnIleAsnGluIle**TyrThr 28
XX      ||||| : : : : :
XX      64 CAATTCATGCAATTGCACCACTCT 87
XX
XX      RESULT 10
XX      ID AAQ73582
XX      AAQ73582 standard; DNA; 561 BP.
XX
XX      AAQ73582;
XX
XX      25-JUN-1995 (first entry)
XX
XX      Human manganese superoxide dismutase gene.
XX
XX      MnSOD; tetramer; pulmonary oxygen toxicity; mitochondrial; mutant;
XX      88.
XX
XX      Homo sapiens.
XX
XX      MO9421283-A.
XX

```

```

XX 29-SEP-1994.
PD
XX
XX 24-MAR-1994; 94WO-US03185.
PF
XX
XX 24-MAR-1993; 93US-0036604.
PR
XX
PA (BOEH ) BOEHRINGER INGELHEIM PHARM INC.
XX
PI Wegner C, Wolyniec W;
XX
XX WPI; 1994-316659/39.
DR

-

US-09-987-190-2 (1-30) x AAQ73582 (1-561)

Alignment Scores:
Pred. No.: 0.000577 Length: 561
Score: 74.00 Matches: 12
Percent Similarity: 78.57% Conservative: 10
Best Local Similarity: 42.86% Mismatches: 6
Query Match: 51.03% Indels: 0
DB: Gaps: 15

QY 1 LysTYrSerLeuProGluLeuAspTYrGluPheSerAlaThrGluProTYrIleSerGly 200
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
DB 4 AAGCACTCTTTCCAGACTTCGACATACGACTACGGTGTCTAGAACCAACACATCAATGCT 633
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
QY 21 GlnIleAsnGluIle**tyrThr 28
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
DB 64 CAAATCATGCAATTCGACCACTCT 87

RESULT 11
AAQ73583
ID AAQ73583 standard; DNA; 561 BP.
XX
AC AAQ73583;
XX
DT 25-JUN-1995 (first entry)
XX
DE Human manganese superoxide dismutase variant gene.
XX
KW MnSOD; tetramer; pulmonary oxygen toxicity; mitochondrial; mutant;
XX ss.
XX
OS Homo sapiens.
XX
PN WO9421283-A.
XX
PD 29-SEP-1994.
XX
PF 24-MAR-1994; 94WO-US03185.
XX
PR 24-MAR-1993; 93US-0036604.
XX
PA (BOEH ) BOEHRINGER INGELHEIM PHARM INC.
XX
PI Wegner C, Wolyniec W;
XX
XX WPI; 1994-316659/39.
DR

```

XX Improved inhibition of pulmonary oxygen toxicity - by  
PT propylactic, topical admin. of human mitochondrial manganese  
PT superoxidismutase  
XX  
PS Claim 7, Page 26; 41pp; English.  
XX  
CC The sequence is that of a mutant human manganese superoxide  
CC dismutase tetramer. Mutations of the native protein provide  
CC a protein which, when applied topically, can inhibit pulmonary  
CC oxygen toxicity.  
CC See also AAQ73582-5.  
XX  
SQ Sequence 561 BP; 159 A; 128 C; 139 G; 135 T; 0 other;  
Alignment Scores:  
Pred. No.: 0.000577 Length: 561  
Score: 74.00 Matches: 12  
Percent Similarity: 78.57% Conservative: 10  
Local Similarity: 42.86% Mismatches: 6  
Query Match: 51.03% Indels: 0  
Gaps: 0  
DB: 15  
US-09-987-190-2 (1-30) x AAQ73583 (1-561)  
QY 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20  
DB 4 AAGCAGCTTTGGCAGACTTGCATACGACTACGCTCTAGAACACACATCATGCT 63  
QY 21 GlnIleAsnGluIle\*\*\*TyrThr 28  
DB 64 CAATCATGCATATGCACACTCT 87  
RESULT 12  
AAQ73584  
ID AAQ73584 standard; DNA; 561 BP.  
XX  
AC AAQ73584;  
XX  
DT 25-JUN-1995 (first entry)  
XX  
DE Human manganese superoxide dismutase variant gene.  
XX  
KM MnSOD; tetramer; pulmonary oxygen toxicity; mitochondrial; mutant;  
KM ss.  
XX  
OS Homo sapiens.  
XX  
PA WO9421283-A.  
XX  
PD 29-SEP-1994.  
XX  
PF 24-MAR-1994; 94WO-US03185.  
XX  
PR 24-MAR-1993; 93US-0036604.  
XX  
PA (BOEH ) BOEHRINGER INGELHEIM PHARM INC.  
XX  
PI Wegner C, Wolyniec WW;  
XX  
DR WPI; 1994-316659/39.  
XX  
XX Improved inhibition of pulmonary oxygen toxicity - by  
PT propylactic, topical admin. of human mitochondrial manganese  
PT superoxidismutase  
XX  
PS Claim 7; Page 27; 41pp; English.  
XX  
CC The sequence is that of a mutant human manganese superoxide  
CC dismutase tetramer. Mutations of the native protein provide  
CC a protein which, when applied topically, can inhibit pulmonary  
CC oxygen toxicity.  
CC See also AAQ73582-5.

XX SQ Sequence 561 BP; 160 A; 126 C; 139 G; 136 T; 0 other;  
Alignment Scores:  
Pred. No.: 0.000577 Length: 561  
Score: 74.00 Matches: 12  
Percent Similarity: 78.57% Conservative: 10  
Best Local Similarity: 42.86% Mismatches: 6  
Query Match: 51.03% Indels: 0  
Gaps: 0  
DB: 15  
US-09-987-190-2 (1-30) x AAQ73584 (1-561)  
QY 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20  
DB 4 AAGCAGCTTTGGCAGACTTGCATACGACTACGCTCTAGAACACACATCATGCT 63  
QY 21 GlnIleAsnGluIle\*\*\*TyrThr 28  
DB 64 CAATCATGCATATGCACACTCT 87  
RESULT 13  
AAZ39779  
ID AAZ39779 standard; cDNA; 594 BP.  
XX  
AC AAZ39779;  
XX  
DT 06-MAR-2000 (first entry)  
XX  
DE Human manganese superoxide dismutase (MnSOD) protein encoding cDNA.  
XX  
KM Manganese superoxide dismutase; MnSOD; ECsOD; proteoglycan; human;  
KM endothelial; oxidative damage; pulmonary inflammatory injury; cancer;  
KM lung disease; hypoxia; ischemia; reperfusion injury; arthritis;  
KM hyperoxia; atherosclerosis; lupus erythematosus; hypertension;  
KM neutrophil-mediated inflammation; ss.  
XX  
OS Homo sapiens.  
XX  
PA WO958547-A1.  
XX  
PD 18-NOV-1999.  
XX  
PF 06-MAY-1999; 99WO-US09921.  
XX  
PR 08-MAY-1998; 98US-0075019.  
XX  
PA (WEBB-) WEBB-WARING INST BIOMEDICAL RES.  
XX  
PI McCord JM, Gao B, Flores SC;  
XX  
DR WPI; 2000-062283/05.  
XX  
DR P-PSDB; AAY55846.  
XX  
PT Modified manganese superoxide dismutase, methods of production and  
PT antibodies -  
XX  
PS Claim 6; Page 71-72; 83pp; English.  
XX  
XX The invention provides a nucleic acid molecule encoding a genetically  
XX modified manganese superoxide dismutase (MnSOD). The nucleic acid  
XX comprises a first nucleic acid sequence encoding an enzymatically active  
XX portion of MnSOD; and a second nucleic acid sequence encoding a peptide  
XX (ECsOD) that binds to polyanionic polysaccharides or proteoglycans on  
XX endothelial cell surfaces. The protein protects a mammal, especially  
XX humans, from oxidative damage and especially from conditions consisting  
XX of pulmonary inflammatory injury, lung disease, cancer, hypoxia, ischemia  
XX reperfusion injury, hyperoxia, atherosclerosis, arthritis, lupus  
XX erythematosus, hypertension and neutrophil-mediated inflammation. The  
XX lung disease is infant or adult respiratory distress syndrome.  
XX interstitial lung disease or asthma. The mutant MnSOD also protects  
XX organs of mammals from pre- and post-transplantation oxidative damage.  
XX MnSOD is positively charged at physiological pH and has a longer plasma

CC half-life of about 4 hours compared to 6-15 minutes for cytosolic Cu,Zn  
 CC SOD (following intravenous injection). ECSD has a substantial advantage  
 CC over both Cu,ZnSOD and MnSOD because of its ability to bind to the  
 CC endothelium. ECSD is hard to produce in mammalian cell culture systems,  
 CC so a fusion of ECSD and MnSOD therefore overcomes these problems. The  
 CC present sequence represents a cDNA encoding a human MnSOD protein.  
 XX  
 XX Sequence 594 BP; 166 A; 144 C; 154 G; 130 T; 0 other;

## Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score:	0.000621	594	12	10	0	0
Percent Similarity:	74.00					
Best Local Similarity:	78.57%					
Query Match:	42.86%					
DB:	51.03%					
	21					

US-09-987-190-2 (1-30) x AA239779 (1-594)

1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20

1 AAGCACAGCTCTCCCGACTGCTCCACACTACGCGCCCTGGAACTCAGATCAACGG 60

21 GlnIleAsnGluIle\*\*TyrThr 28

61 CAGATCATGACAGCTGCACACAGC 84

RESULT 14

AA81219 standard; cDNA; 600 BP.

AA81219;

03-OCT-1990 (first entry)

Sequence of Formula IIb encoding modified portion of human manganese

superoxide dismutase (hMnSOD).

Human manganese superoxide dismutase derivative; clone B8;

formula IIb; enzyme; EC-1.15.1.1.

Homo sapiens.

Key CDS

Location/Qualifiers

4..600

/\*tag= a

EP282899-A.

21-SEP-1988.

10-MAR-1988; 88EP-0103754.

24-DEC-1987; 87DE-3744038.

14-MAR-1987; 87DE-3708306.

(BOEH) BOEHRINGER INGELHEIM.

Heckl K, Spevak W, Ostermann E, Zophel A, Krystek E, Maurer-Fogy I;

Whiche-Castanon MJ, Stralowa C, Hauptmann R;

WPI; 1988-265361/38.

P-PSDB; AAP80603.

Prodn. of human manganese superoxidizedismutase peptide(s) -

and DNA coding sequences, for control and diagnosis of eg

inflammatory diseases

CC encoding all or part of the polypeptides and replicating vectors,  
 CC expression plasmids and transformed host cells contg. these sequences.  
 CC hMnSOD catalyses disproportionation of the superoxide radical and can be  
 CC used in the prevention, diagnosis and treatment of inflammatory,  
 CC degenerative, neoplastic and rheumatoid disorders; for wound healing; in  
 CC autoimmune disease and in organ transplantation. It can also be used to  
 CC improve the storage stability of liq. and solid foods. AA81219 is  
 CC derived from Formula Ib (AA81826).  
 XX  
 XX Sequence 600 BP; 172 A; 135 C; 150 G; 143 T; 0 other;

## Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score:	0.000629	600	12	10	0	0
Percent Similarity:	74.00					
Best Local Similarity:	78.57%					
Query Match:	42.86%					
DB:	51.03%					
	9					

US-09-987-190-2 (1-30) x AA81219 (1-600)

1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20

4 AAGCACTCTTGCAGACTGCTCCACACTACGCGCTCTAGAACACACATCAATGCT 63

21 GlnIleAsnGluIle\*\*TyrThr 28

64 CAATCATGCAATTCACACACTCT 87

RESULT 15

AA81218 standard; cDNA; 600 BP.

AA81218;

03-OCT-1990 (first entry)

Sequence of Formula IIA encoding modified portion of human manganese

superoxide dismutase (hMnSOD).

Human manganese superoxide dismutase derivative; clone B8;

formula IIA; enzyme; EC-1.15.1.1.

Homo sapiens.

Key CDS

Location/Qualifiers

4..600

/\*tag= a

EP282899-A.

21-SEP-1988.

10-MAR-1988; 88EP-0103754.

24-DEC-1987; 87DE-3744038.

14-MAR-1987; 87DE-3708306.

(BOEH) BOEHRINGER INGELHEIM.

Heckl K, Spevak W, Ostermann E, Zophel A, Krystek E, Maurer-Fogy I;

Whiche-Castanon MJ, Stralowa C, Hauptmann R;

WPI; 1988-265361/38.

P-PSDB; AAP80602.

Prodn. of human manganese superoxidizedismutase peptide(s) -

and DNA coding sequences, for control and diagnosis of eg

inflammatory diseases

Disclosure; ; pp; German.

The patent is for polypeptides, esp. non-glycosylated, having the



CC hMnSOD catalyzes disproportionation of the superoxide radical and can be used in the prevention, diagnosis and treatment of inflammatory, CC degenerative, neoplastic and rheumatoid disorders; for wound healing; in CC autoimmune disease and in organ transplantation. It can also be used to CC improve the storage stability of liq. and solid foods. Oligo Via CC (XhoI/XbaI) (AA081222) was ligated to oligo VIB (XbaI/NcoI) (AA081223) CC and then used to produce cDNA encoding the entire hMnSOD of formula VIIA CC (AA081224) and formula VIIB.

XX SQ Sequence 600 BP; 172 A; 134 C; 150 G; 144 T; 0 other;

# Alignment Scores:

Pred. No.:	0.000629	Length:	600
Score:	74.00	Matches:	12
Percent Similarity:	78.57%	Conservative:	10
Best Local Similarity:	42.86%	Mismatches:	6
Query Match:	51.03%	Indels:	0
DB:	9	Gaps:	0

US-09-987-190-2 (1-30) x AA081225 (1-600)

1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20  
 4 AAGCAGCTCTTGCCAGACTTGCATACGACTACGCTGCTTGAACACACATCATGCT 63

Qy 21 GlnIleAngIuile\*\*TyrThr 28  
 64 CAATCATGCAATGACACACTCT 87

# RESULT 18

AA094271 standard; cDNA; 600 BP.

AC AA094271;

DT 17-MAY-1996 (first entry)

XX Human manganese superoxide dismutase (Lys29) coding sequence.

DE hMn-SOD; manganese; superoxide dismutase; tetramer; inflammation;

KW autoimmune disease; rheumatoid arthritis; neoplasm; emphysema;

KW osteoarthritis; wound healing; ss.

OS Homo sapiens.

XX Location/Qualifiers  
 FT CDS 1..600  
 /\*tag= a  
 /product= Met\_hMn-SOD (Lys29)

EP676472-A1.

XX 11-OCT-1995.

XX 10-MAR-1988; 88EP-0107460.

XX 24-DEC-1987; 87DE-3744038.

XX 14-MAR-1987; 87DE-3708306.

XX 26-MAY-1987; 87DE-3717695.

XX 10-JUL-1987; 87DE-3722884.

XX (BOEH) BOEHRINGER INGELHEIM INT GMBH.

XX Hauptmann R, Heckl K, Krystek E, Maurer-Fogy I;  
 PI Ostermann E, Spevak W, Stratawa C, Wiche-Castanon M;  
 PI Zephel A;

XX WPI; 1995-346092/45.

XX P-PSDB; AAR75191.

XX Genes encoding recombinant human manganese superoxide dismutase -  
 PT for treatment, prevention and diagnosis of inflammatory diseases  
 XX

PS Claim 8; Page 6; 54pp; German.

XX The present sequence codes for a Met residue followed by the 198  
 CC amino acid long hMn-SOD mature polypeptide having a Lys residue  
 CC at position 29. The cDNA can be inserted into expression vectors  
 CC containing appropriate signal sequences, etc. for expression of  
 CC recombinant hMn-SOD. The N-terminal Met residue can be removed,  
 CC e.g. by treatment with CNBr or CNCl. Mature hMn-SOD is useful for  
 CC treating, preventing or diagnosing inflammatory, degenerative,  
 CC neoplastic and rheumatic disorders; in wound healing and in  
 CC autoimmune diseases.

XX SQ Sequence 600 BP; 173 A; 134 C; 150 G; 143 T; 0 other;

# Alignment Scores:

Pred. No.:	0.000629	Length:	600
Score:	74.00	Matches:	12
Percent Similarity:	78.57%	Conservative:	10
Best Local Similarity:	42.86%	Mismatches:	6
Query Match:	51.03%	Indels:	0
DB:	16	Gaps:	0

US-09-987-190-2 (1-30) x AA094271 (1-600)

Qy 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20  
 4 AAGCAGCTCTTGCCAGACTTGCATACGACTACGCTGCTTGAACACACATCATGCT 63

Qy 21 GlnIleAngIuile\*\*TyrThr 28  
 64 CAATCATGCAATGACACACTCT 87

# RESULT 19

AA094272 standard; cDNA; 600 BP.

AC AA094272;

DT 17-MAY-1996 (first entry)

XX Human manganese superoxide dismutase (Gln29) coding sequence.

DE hMn-SOD; manganese; superoxide dismutase; tetramer; inflammation;

KW autoimmune disease; rheumatoid arthritis; neoplasm; emphysema;

KW osteoarthritis; wound healing; ss.

OS Homo sapiens.

XX Location/Qualifiers  
 FT CDS 1..600  
 /\*tag= a  
 /product= Met\_hMn-SOD (Gln29)

EP676472-A1.

XX 11-OCT-1995.

XX 10-MAR-1988; 88EP-0107460.

XX 24-DEC-1987; 87DE-3744038.

XX 14-MAR-1987; 87DE-3708306.

XX 26-MAY-1987; 87DE-3717695.

XX 10-JUL-1987; 87DE-3722884.

XX (BOEH) BOEHRINGER INGELHEIM INT GMBH.

XX Hauptmann R, Heckl K, Krystek E, Maurer-Fogy I;  
 PI Ostermann E, Spevak W, Stratawa C, Wiche-Castanon M;  
 PI Zephel A;

XX WPI; 1995-346092/45.

XX P-PSDB; AAR75192.



PT Genes encoding recombinant human manganese superoxide dismutase -  
 PT for treatment, prevention and diagnosis of inflammatory diseases  
 PS Claim 8; Page 6; 54pp; German.  
 XX  
 CC The present sequence codes for a Met residue followed by the 198  
 CC amino acid long hMn-SOD mature polypeptide having a Gln residue  
 CC at position 29. The cDNA can be inserted into expression vectors  
 CC containing appropriate signal sequences, etc. for expression of  
 CC recombinant hMn-SOD. The N-terminal Met residue can be removed.  
 CC e.g. by treatment with CNBr or CNCl. Mature hMn-SOD is useful for  
 CC treating, preventing or diagnosing inflammatory, degenerative,  
 CC neoplastic and rheumatic disorders; in wound healing and in  
 CC autoimmune diseases.  
 CC  
 SQ Sequence 600 BP; 172 A; 135 C; 150 G; 143 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 0.000629 Length: 600  
 Score: 74.00 Matches: 12  
 Percent Similarity: 78.57% Conservative: 10  
 Best Local Similarity: 42.86% Mismatches: 6  
 Query Match: 51.03% Indels: 0  
 DB: 16 Gaps: 0  
 US-09-987-190-2 (1-30) x AAQ94272 (1-600)  
 QY 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20  
 Db 4 AAGCAGCTCTTTGCCAGACTTCCCATACGACTACGCTCTCTAGAACACACATCAATGCT 63  
 QY 21 GlnIleAsnGluIle\*\*TyrThr 28  
 Db 64 CAATCATGCATATGCACCACTCT 87  
 RESULT 20  
 AAQ94277 standard; cDNA; 600 BP.  
 AC AAQ94277;  
 XX  
 DT 23-MAY-1996 (first entry)  
 XX  
 DE Human manganese superoxide dismutase (Lys29) coding sequence.  
 XX  
 KM hMn-SOD; manganese; superoxide dismutase; tetramer; inflammation;  
 KM autoimmune disease; rheumatoid arthritis; neoplasm; emphysema;  
 KM osteoarthritis; wound healing; ss.  
 OS Homo sapiens.  
 OS Synthetic.  
 FT Key location/Qualifiers  
 FT CDS 1..600  
 FT /\*tag= a  
 FT /product= Met hMn-SOD(Lys29)  
 FT /note= "the first 31 codons are provided by a  
 FT synthetic linker sequence"  
 XX  
 PN EP676472-A1.  
 PD 11-OCT-1995.  
 XX  
 PF 10-MAR-1988; 88EP-0107460.  
 XX  
 PR 24-DEC-1987; 87DE-3744038.  
 PR 14-MAR-1987; 87DE-3708306.  
 PR 26-MAY-1987; 87DE-3717695.  
 PR 10-JUL-1987; 87DE-3722884.  
 XX  
 PA (BOEH ) BOEHRINGER INGELHEIM INT GMBH.  
 PI Hauptmann R, Heckl K, Krystek E, Maurer-Fogy I;

PI Ostermann E, Spevak W, Stralowa C, Wiche-Castanon M;  
 PI Zepherl A;  
 XX  
 DR WPI; 1995-346092/45.  
 XX  
 PT Genes encoding recombinant human manganese superoxide dismutase -  
 PT for treatment, prevention and diagnosis of inflammatory diseases  
 PS Claim 8; Page 17; 54pp; German.  
 XX  
 CC The present sequence codes for a Met residue followed by the 198  
 CC amino acid long hMn-SOD mature polypeptide having a Lys residue  
 CC at position 29. The cDNA can be inserted into expression vectors  
 CC containing appropriate signal sequences, etc. for expression of  
 CC recombinant hMn-SOD. The N-terminal Met residue can be removed.  
 CC e.g. by treatment with CNBr or CNCl. Mature hMn-SOD is useful for  
 CC treating, preventing or diagnosing inflammatory, degenerative,  
 CC neoplastic and rheumatic disorders; in wound healing and in  
 CC autoimmune diseases.  
 CC  
 SQ Sequence 600 BP; 173 A; 133 C; 150 G; 144 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 0.000629 Length: 600  
 Score: 74.00 Matches: 12  
 Percent Similarity: 78.57% Conservative: 10  
 Best Local Similarity: 42.86% Mismatches: 6  
 Query Match: 51.03% Indels: 0  
 DB: 16 Gaps: 0  
 US-09-987-190-2 (1-30) x AAQ94277 (1-600)  
 QY 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20  
 Db 4 AAGCAGCTCTTTGCCAGACTTCCCATACGACTACGCTCTCTAGAACACACATCAATGCT 63  
 QY 21 GlnIleAsnGluIle\*\*TyrThr 28  
 Db 64 CAATCATGCATATGCACCACTCT 87  
 RESULT 21  
 AAQ94278 standard; cDNA; 600 BP.  
 AC AAQ94278;  
 XX  
 DT 23-MAY-1996 (first entry)  
 XX  
 DE Human manganese superoxide dismutase (Gln29) coding sequence.  
 XX  
 KM hMn-SOD; manganese; superoxide dismutase; tetramer; inflammation;  
 KM autoimmune disease; rheumatoid arthritis; neoplasm; emphysema;  
 KM osteoarthritis; wound healing; ss.  
 OS Homo sapiens.  
 OS Synthetic.  
 FT Key location/Qualifiers  
 FT CDS 1..600  
 FT /\*tag= a  
 FT /product= Met hMn-SOD(Gln29)  
 FT /note= "the first 31 codons are provided by a  
 FT synthetic linker sequence"  
 XX  
 PN EP676472-A1.  
 PD 11-OCT-1995.  
 XX  
 PF 10-MAR-1988; 88EP-0107460.  
 XX  
 PR 24-DEC-1987; 87DE-3744038.  
 PR 14-MAR-1987; 87DE-3708306.  
 PR 26-MAY-1987; 87DE-3717695.

PR 10-JUL-1987; 87DE-3722884.  
 XX  
 PA (BOEH ) BOEHRINGER INGELHEIM INT GMBH.  
 XX  
 PI Hauptmann R, Heckl K, Krystek E, Maurer-Fogy I,  
 PI Oestermann E, Spevak W, Stralow C, Wiche-Caetanon M,  
 PI Zoeschel A;  
 XX  
 DR WPI; 1995-346092/45.  
 XX  
 PT Genes encoding recombinant human manganese superoxide dismutase -  
 PT for treatment, prevention and diagnosis of inflammatory diseases  
 XX  
 PS Claim 8; Page 17-18; 54pp; German.  
 XX  
 CC The present sequence codes for a Met residue followed by the 198  
 CC amino acid long hm-SOD mature polypeptide having a Gln residue  
 CC at position 29. The CDNA can be inserted into expression vectors  
 CC containing appropriate signal sequences, etc. for expression of  
 CC recombinant hm-SOD. The N-terminal Met residue can be removed,  
 CC e.g. by treatment with CNBr or CNCL. Mature hm-SOD is useful for  
 CC treating, preventing or diagnosing inflammatory, degenerative,  
 CC neoplastic and rheumatic disorders; in wound healing and in  
 CC autoimmune diseases.  
 CC  
 SQ Sequence 600 BP; 172 A; 134 C; 150 G; 144 T; 0 other;  
 XX  
 Alignment Scores:  
 Pred. No.: 0.000629 Length: 600  
 Score: 74.00 Matches: 12  
 Percent Similarity: 78.57% Conservative: 10  
 Best Local Similarity: 42.86% Mismatches: 6  
 Query Match: 51.03% Indels: 0  
 DB: 16 Gaps: 0  
 US-09-987-190-2 (1-30) x AAQ20194 (1-600)  
 QY 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSergly 20  
 Db 4 AAGCAGCTTTGCCGAGCTTCCGACACGACGCTGCTCTGAAACACACATCATCT 63  
 QY 21 GlnIleAsngIuile\*\*TyrThr 28  
 Db 64 CAATCATGCAATGCACCATCT 87  
 RESULT 22  
 AAQ20194  
 ID AAQ20194 standard; DNA; 609 BP.  
 AAQ20194;  
 DT 01-APR-1992 (first entry)  
 XX  
 DE Mn-SOD (Ile59Thr).  
 XX  
 KM Manganese superoxide dismutase; vector; anti-inflammatory; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EP462836-A.  
 PD 27-DEC-1991.  
 PF 20-JUN-1991; 91EP-0305596.  
 PR 19-OCT-1990; 90JP-0279286.  
 PR 20-JUN-1990; 90JP-0159925.  
 XX  
 PA (MITK ) MITSUI TOATSU CHEM INC.  
 PI Takahashi S, Makino T, Asanagi M, Yoshino C;  
 XX  
 DR WPI; 1992-001187/01.

DR P-PSDB; AAR20015.  
 XX  
 PT New recombinant vector plasmid - expresses human manganese,  
 PT superoxide dismutase, used as antiinflammatory agent  
 XX  
 PS Disclosure; Fig 8; 22pp; English.  
 XX  
 CC A vector was constructed contg. lac/P1 ligated promoter which has  
 CC this human Mn-SOD structural gene. The gene encodes modified human  
 CC Mn-SOD in which the 59th isoleucine is converted to threonine.  
 XX  
 SQ Sequence 609 BP; 171 A; 148 C; 157 G; 133 T; 0 other;  
 XX  
 Alignment Scores:  
 Pred. No.: 0.000641 Length: 609  
 Score: 74.00 Matches: 12  
 Percent Similarity: 78.57% Conservative: 10  
 Best Local Similarity: 42.86% Mismatches: 6  
 Query Match: 51.03% Indels: 0  
 DB: 13 Gaps: 0  
 US-09-987-190-2 (1-30) x AAQ20194 (1-609)  
 QY 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSergly 20  
 Db 6 AAGCAGAGCCCTCCCGACCTGCCCTACGACACGCGCCCTGGAACCTCATCATCAGCG 65  
 QY 21 GlnIleAsngIuile\*\*TyrThr 28  
 Db 66 CAGATCATGCGAGCTGCACCATCAGC 89  
 RESULT 23  
 AA239781  
 ID AA239781 standard; DNA; 681 BP.  
 XX  
 AC AA239781;  
 XX  
 DT 06-MAR-2000 (first entry)  
 XX  
 DE Human manganese superoxide dismutase (MnSOD) mutant protein DNA.  
 XX  
 KM Manganese superoxide dismutase; MnSOD; EC503; proteoglycan; human;  
 KM endotelial; oxidative damage; pulmonary inflammatory injury; cancer;  
 KM lung disease; hypoxia; ischemia; reperfusion injury; arthritis;  
 KM hyperoxia; atherosclerosis; lupus erythematosus; hypertension;  
 KM neutrophil-mediated inflammation; mutant; ss.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 FN WO9558547-A1.  
 PD 18-NOV-1999.  
 PF 06-MAY-1999; 99WO-US09921.  
 PR 08-MAY-1998; 98US-0075019.  
 XX  
 PA (WEBB-) WEBB-WARING INST BIOMEDICAL RES.  
 PI McCord JM, Gao B, Flores SC;  
 XX  
 DR WPI; 2000-062283/05.  
 DR P-PSDB; AAY55849.  
 XX  
 PT Modified manganese superoxide dismutase, methods of production and  
 PT antibodies -  
 XX  
 PS Claim 11; Page 74-75; 83pp; English.  
 CC The invention provides a nucleic acid molecule encoding a genetically  
 CC modified manganese superoxide dismutase (MnSOD). The nucleic acid  
 CC comprises a first nucleic acid sequence encoding an enzymatically active

CC portion of MsOD: and a second nucleic acid sequence encoding a peptide  
CC (ECSD) that binds to polyanionic polysaccharides or proteoglycans on  
CC endothelial cell surfaces. The protein protects a mammal, especially  
CC humans, from oxidative damage and especially from conditions consisting  
CC of pulmonary inflammatory injury, lung disease, cancer, hypoxia, ischemia  
CC reperfusion injury, hyperoxia, atherosclerosis, arthritis, lupus  
CC erythematosus, hypertension and neutrophil-mediated inflammation. The  
CC lung disease is infant or adult respiratory distress syndrome. The  
CC interstitial lung disease or asthma. The mutant MsOD also protects  
CC organs of mammals from pre- and post-transplantation oxidative damage.  
CC MsOD is positively charged at physiological pH and has a longer plasma  
CC half-life of about 4 hours compared to 6-15 minutes for cytosolic Cu,Zn  
CC SOD (following intravenous injection). ECSD has a substantial advantage  
CC over both Cu,ZnSOD and MsOD because of its ability to bind to the  
CC endothelium. ECSD is hard to produce in mammalian cell culture systems,  
CC so a fusion of ECSD and MsOD therefore overcomes these problems. The  
CC present sequence represents a DNA encoding a human MsOD mutant protein.  
CC This comprises the MsOD sequence, and a 78 basepair ECSD tail-encoding  
CC fragment.

Sequence 681 BP; 185 A; 170 C; 189 G; 137 T; 0 other;

**Alignment Scores:**

Pred. No.:	0.00074	Length:	681
Score:	74.00	Matches:	12
Percent Similarity:	78.57%	Conservative:	10
Best Local Similarity:	42.86%	Mismatches:	6
Query Match:	51.03%	Indels:	0
DB:	21	Gaps:	0

US-09-987-190-2 (1-30) x AAZ39781 (1-681)

```
Oy      1 LysTyrSerLeuProGlnLeuAspTyrGlnIleHisSerAlaThrGluProTyrIleSergly 20  
       |||:::|||||:::|||::|::|::|::|::|::|::|::|::|::|::|::|:  
Db      7 AAGCACAGCCCTCCCCGACTGCCTCACGATCAACGGCGGCCCTGTGAACCTCATCATCAACGCG 66
```

```

Qy      21 GlnIleAsnGluIle**TyrThr 28
          |||||
Db      67 CAGATCATGCAGCTGCACACACAGC 90

```

RESULT 24  
AAN71370  
ID AAN71370 standard; DNA; 813 BP.

AC AAN71370 ;

DT 14-MAY-1991 (first entry)

Sequence encoding human manganese superoxide dismutase.

KW Human MSOD; hydrogen peroxide; ischaemia; lesions; inflammation; free radicals; ds.

FH	Key	location/Qualifiers
FT	CDS	43..711
FT		/*tag= a
FT	mat_peptide	115..708
FT		/*tag= b
FT		/product= mature MSOD

PN BE905796-A.

PD 16-MAR-1987.

PF 20-NOV-1986; 86BE-0905796.

PR	12-SEP-1986;	86US-0907051.
PR	22-NOV-1985;	85US-0801090.

PA (BIOT-) BIO-TECHNOL GEN.

DR WPI; 1987-101441/15.  
DR P-PSDB; AAP71370.

XX New DNA coding for polypeptide of human manganese superoxide-  
PT dismutase - useful e.g. for treating inflammation, and new  
PT expression vectors and transformed cells  
XX  
PS Disclosure, fig 1; 46pp; French.

PS Disclosure; fig 1; 46pp; French.

CC This purified cDNA encodes bot. prepro- and mature-manganese superoxide dismutase (MSOD). It is one strand of a double stranded CC molecule contained in a recombinant vehicle. The MSOD produced CC catalyses the reaction of hydrogen ions and the SOD radical anion to form hydrogen peroxide and water. It is useful in veterinary and CC pharmaceutical compans. for e.g. reducing lesions of reperfusion CC following ischaemia, to prolong survival time isolated organs and CC for treating inflammation. See also AAN71371-72.

**SQ** Sequence 813 BP; 218 A; 204 C; 213 G; 178 T; 0 other;

**Alignment Scores:**

Pred. No.:	0.00928	813
Score:	74.00	Matches: 12
Percent Similarity:	78.57%	Conservative: 10
Best Local Similarity:	42.86%	Mismatches: 6
Query Match:	51.03%	Indels: 0
DB:	8	Gaps: 0

US-09-987-190-2 (1-30) X AAN71370 (1-813)

Qy 1 LysTyrSerLeuProGluLeuAspTyrGluPhenSerAlaThrGluProTyrIleSercyl 20  
|||:::|||||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||  
Db 115 AAGCACAGCTCCCCGACTGCCTACGACTACGGCGCCCTGGAACCTCACATCAACCG 174

```
Qy      21 GlnIleAsnGluIle**TyrThr 28
          |||||
Db      175 CAGATCATGCAGCTGCACCAACAGC 198
```

RESULT 25  
AAN81158  
ID AAN81158 standard; cDNA; 813 BP.

AC PAN81158;

DT 22-OCT-1990 (first entry)

DE cDNA encoding human manganese superoxide dismutase

KM Human manganese superoxide dismutase; ds cDNA; oxygen free radicals.

Os Homo sapiens.

FH	Key	Location/Qualifiers
FT	CDS	43..711
FT		/*tag= a

PN EP284105-A.

PD 28-SEP-1988

PF 25-MAR-1988; 88EP-0104880.

PR 26-FEB-1988; 88US-0161117  
PR 27-MAR-1987; 87US-0032734.

PA (BIOT-) BIO-TECHNOLOGY GENERAL CORPORATION

PI Hartman JR, Beck Y, Nimrod A;

DR WPI; 1988-272584/39  
DR P-PSDB; AAP80551.

PT Recombinant human manganese superoxidizedismutase -  
PT used for treating, eg reperfusion injury, inflammation,  
PT arthritis, bronchial pulmonary displasia or lung fibrosis

XX Disclosure; ; pp: English.  
XX The cDNA encodes human manganese superoxide dismutase, and is  
CC inserted into a plasmid, eg pMSB-4 (ATCC 53250).  
XX  
SQ Sequence 813 BP; 218 A; 203 C; 214 G; 178 T; 0 other;  
Alignment Scores:  
Pred. No.: 0.000928 Length: 813  
Score: 74.00 Matches: 12  
Percent Similarity: 78.57% Conservative: 10  
Best Local Similarity: 42.86% Mismatches: 6  
Query Match: 51.03% Indels: 0  
Gaps: 0  
DB: 9  
US-09-987-190-2 (1-30) x AAN81158 (1-813)  
Oy 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20  
115 AAGCAGACGCTCCCGACCTGCTACGACTACGCGCGCTCGAAGCTCAGATCAACGCG 174  
Oy 21 GlnIleAsnGluIle\*\*\*TyrThr 28  
115 AAGCAGACGCTCCCGACCTGCTACGACTACGCGCGCTCGAAGCTCAGATCAACGCG 174  
Db 175 CAGATCATGACGCTGCACACACGAC 198  
RESULT 26  
AAQ53193  
ID AAQ53193 standard; cDNA; 813 BP.  
XX AAQ53193;  
AC  
XX 21-JUN-1994 (first entry)  
DT  
XX  
DE MnsOD cDNA.  
XX  
XX MnsOD; manganese superoxide dismutase; N-terminal; catalytic;  
KM reperfusion injury; ischaemia; superoxide; SO; molecular oxygen;  
KM anti-inflammatory; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 43..711  
FT /\*tag= a  
FT /product= MnsOD  
XX  
XX US5270195-A.  
XX  
XX 14-DEC-1993.  
XX  
XX 22-NOV-1985; 85US-0801090.  
XX  
XX 29-OCT-1986; 86IE-0002851.  
XX  
XX (BIOT-) BIO-TECHNOLOGY GENERAL CORP.  
XX  
XX Beck Y, Hartman JR;  
XX  
XX WPI; 1993-404931/50.  
XX  
XX P-PSDB; AAR44801.  
XX  
XX Expression plasmid in Escherichia coli host system - encodes  
PT human manganese superoxidodismutase analogue, useful for e.g.  
PT creating inflammation  
XX  
XX  
XX Claim 1; Fig 1a-1c; 27pp; English.  
XX  
XX The sequence encodes a manganese superoxide dismutase which can be used  
CC to catalyse the reduction of superoxide (SO) radicals to hydrogen  
CC peroxide and molecular oxygen. It can be used to reduce  
CC reperfusion injury following ischaemia and prolong the survival of  
CC excised organs. It can also be used as a long acting

CC anti-inflammatory drug.  
XX  
XX Sequence 813 BP; 218 A; 203 C; 214 G; 178 T; 0 other;  
SQ  
Alignment Scores:  
Pred. No.: 0.000928 Length: 813  
Score: 74.00 Matches: 12  
Percent Similarity: 78.57% Conservative: 10  
Best Local Similarity: 42.86% Mismatches: 6  
Query Match: 51.03% Indels: 0  
Gaps: 0  
DB: 14  
US-09-987-190-2 (1-30) x AAQ53193 (1-813)  
Oy 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20  
115 AAGCAGACGCTCCCGACCTGCTACGACTACGCGCGCTCGAAGCTCAGATCAACGCG 174  
Db 115 AAGCAGACGCTCCCGACCTGCTACGACTACGCGCGCTCGAAGCTCAGATCAACGCG 174  
Oy 21 GlnIleAsnGluIle\*\*\*TyrThr 28  
115 AAGCAGACGCTCCCGACCTGCTACGACTACGCGCGCTCGAAGCTCAGATCAACGCG 174  
Db 175 CAGATCATGACGCTGCACACACGAC 198  
RESULT 27  
AAT34277  
ID AAT34277 standard; cDNA; 813 BP.  
XX AAT34277;  
AC  
XX 24-OCT-1996 (first entry)  
DT  
XX  
DE Human manganese superoxide dismutase cDNA.  
XX  
XX Manganese superoxide dismutase; MnsOD; reperfusion injury;  
KM ischaemia; bronchial pulmonary dysplasia; inflammation;  
KM antiinflammatory; ds.  
XX  
XX Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 43..711  
FT /\*tag= a  
FT sig\_peptide 43..114  
FT /\*tag= b  
FT mat\_peptide 115..708  
FT /\*tag= c  
XX  
XX US5540911-A.  
XX  
XX 30-JUL-1996.  
XX  
XX 22-NOV-1985; 85US-0801090.  
XX  
XX 29-OCT-1986; 86IE-0002851.  
XX  
XX (BIOT-) BIO-TECHNOLOGY GENERAL CORP.  
XX  
XX Beck Y, Hartman JR;  
XX  
XX WPI; 1996-361912/36.  
XX  
XX P-PSDB; AAM00018.  
XX  
XX Use of recombinant human manganese superoxide dismutase - for  
PT treating inflammation or bronchial pulmonary dysplasia, reducing  
PT reperfusion injury or prolonging organ survival  
XX  
XX  
XX Claim 1; Fig 1A-1C; 27pp; English.  
XX  
XX A cDNA clone (AAT34277) codes for human manganese superoxide dismutase  
CC (MnsOD) (AAM00018), an enzyme that catalyses the reduction of  
CC superoxide radicals to H2O2 and O2. It was isolated from a human T-  
CC cell library cloned into vector gcl0 using probes (see also AAT34283-  
CC 84) based on portions of mature MnsOD. The MnsOD gene was also  
CC identified (see also AAT34278-82). The cDNA can be used for the

CC prodn. of recombinant MnSOD using e.g. Escherichia coli cells as  
 CC hosts. The MnSOD is useful for reducing reperfusion injury,  
 CC treating inflammation or bronchial pulmonary dysplasia and for  
 CC prolonging survival of excised organs.

XX Sequence 813 BP; 218 A; 203 C; 214 G; 178 T; 0 other;

#### Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score: 74.00	813	12	10	6	0	0
Percent Similarity: 78.57%						
Best Local Similarity: 42.86%						
Query Match: 51.03%						

US-09-987-190-2 (1-30) x AAT34277 (1-813)

Qy 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20

115 AAGCAGAGCCCTCCCGACCTGCTCCTACGACTACGCGCCCTCGAAGCTCAGATCAACGCG 174

Qy 21 GlnIleAsnGluIle\*\*TyrThr 28

Db 175 CAGATCATGCGAGCTGCACACAGC 198

#### RESULT 28

AAT15589

AC AAT15589;

DT 06-APR-1996 (first entry)

DE Human manganese superoxide-dismutase cDNA.

XX Human; manganese superoxide-dismutase; T-lymphocyte; probe; cDNA;  
 KW Escherichia coli; plasmid pMS-4; N-terminal truncation; cloning;  
 KM aminopeptidase; antiinflammatory; oxygen free radical scavenger;  
 KM synovial inflammation; arthritis; lung fibrosis; ds.

OS Homo sapiens.

Key Location/Qualifiers

FT CDS 43..711

FT sig\_peptide /tag= a

FT mat\_peptide /product= Manganese superoxide-dismutase

FT /note= "EC-1.15.1.1"

FT /tag= b

FT /tag= c

FT /tag= c

FT /tag= c

FT /tag= c

FT /tag= c

FT /tag= c

FT /tag= c

FT /tag= c

FT /tag= c

FT /tag= c

FT /tag= c

FT /tag= c

XX The sequence encodes a human manganese superoxide-dismutase

CC (MnSOD), and has been isolated from a human T-lymphocyte cDNA

CC library in phage lambda-gt10 in Escherichia coli. The sequence of the

CC a 5'-probe (AAT15591) and a 3'-probe (AAT15592). The sequence of the

CC insert in plasmid pMS-4 (obtained by subcloning in plasmid pBR322)

CC is shown. The cDNA may be expressed in e.g. E. coli for production

CC of recombinant MnSOD. The MnSOD product may be cleaved with

CC Aeromonas proteolytica aminopeptidase to produce an N-terminally

CC truncated analogue with lysine and optionally histidine residues

CC removed. The MnSOD analogue may be used in therapy of conditions

CC associated with generation of oxygen free radicals, particularly

CC synovial inflammation, arthritis and lung fibrosis (claimed).

XX Sequence 813 BP; 218 A; 203 C; 214 G; 178 T; 0 other;

Qy 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20

Db 115 AAGCAGAGCCCTCCCGACCTGCTCCTACGACTACGCGCCCTCGAAGCTCAGATCAACGCG 174

Qy 21 GlnIleAsnGluIle\*\*TyrThr 28

Db 175 CAGATCATGCGAGCTGCACACAGC 198

#### RESULT 29

AAX24998

AC AAX24998;

DT 05-JUL-1999 (first entry)

DE Human native manganese superoxide dismutase cDNA.

KW Superoxide dismutase; MnSOD; SOD; human; protein engineering;

KM enzyme engineering; antioxidant; antiinflammatory; cytotoxicity;

KM inflammation; reperfusion injury; therapy; ss.

OS Homo sapiens.

Key Location/Qualifiers

FT CDS 43..711

FT sig\_peptide /tag= a

FT mat\_peptide /tag= b

FT /tag= c

FT /tag= c

FT /tag= c

FT /tag= c

FT /tag= c

FT /tag= c

FT /tag= c

FT /tag= c

FT /tag= c

PT New recombinant human manganese superoxide dismutase proteins  
XX  
PS Disclosure; Page 52-53; 61pp; English.  
XX  
CC This cDNA sequence codes for human manganese superoxide dismutase  
CC (hMnSOD, see AAM98169). Novel MnSOD proteins having catalytic  
CC activity which differs from this natural hMnSOD are claimed. The  
CC modified proteins exhibit reduced or no product inhibition, or have  
CC greater activity, or both, compared to natural hMnSOD. The  
CC modifications involve one or amino acid substitutions within  
CC the active site of the enzyme, especially at residues His-26,  
CC His-30, Tyr-74, His-74, Asp-159, Trp-161 and His-163  
CC (see also AAM98171-77). Nucleic acids encoding the modified hMnSOD  
CC proteins are also claimed. The modified hMnSOD proteins, or  
CC expression vectors in which modified hMnSOD nucleic acid is linked  
CC to a promoter (preferably mammalian), can be used to protect a cell  
CC line from damage caused by superoxide radicals (claimed). They can  
CC also be used to treat subjects suffering from, or at risk of,  
CC cytotoxicity caused by superoxide radicals (claimed). As such,  
CC they can be used as antioxidants in the treatment of a variety of  
CC disorders, including inflammation (claimed), reperfusion injury  
CC following ischemia (claimed), and cellular damage caused by  
CC chemotherapeutic agents.  
XX  
SQ Sequence 813 BP; 218 A; 203 C; 214 G; 178 T; 0 other;  
Alignment Scores:  
Pred. No.: 0.000928 Length: 813  
Score: 74.00 Matches: 12  
Percent Similarity: 78.57% Conservative: 10  
Best Local Similarity: 42.86% Mismatches: 6  
Query Match: 51.03% Indels: 0  
DB: 20 Gaps: 0  
US-09-987-190-2 (1-30) x AAM24998 (1-813)  
QY 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20  
Db 115 AAGCAGACGCTCCCGACCTGCGCTACGACGCGCCCTGGAACCTCAACGCG 174  
QY 21 GlnIleAsnGluIle\*\*\*TyrThr 28  
Db 175 CAGATCATGCAGCTGCACACGAC 198  
RESULT 30  
AAA63891  
ID AAA63891 standard; cDNA; 813 BP.  
XX  
XX AAA63891;  
04-DEC-2000 (first entry)  
XX  
DE cDNA encoding a human manganese superoxide dismutase.  
XX  
KW Human; manganese superoxide dismutase; hMn SOD; superoxide radical;  
KW superoxide radical damage; cytotoxicity; inflammation; ischemia;  
KW reperfusion injury; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 43..711  
FT /tag= a  
FT /transl\_except= (pos: 505..507, aa: Xaa)  
FT /transl\_except= (pos: 511..513, aa: Ala)  
FT /product= "manganese superoxide dismutase"  
FT /note= "Xaa is Gln or Glu"  
FT 43..114  
FT /tag= b  
FT mat\_peptide 115..708  
FT /tag= c  
PN US6107070-A.

XX  
XX 22-AUG-2000.  
XX  
PD 10-SEP-1998; 98US-0151052.  
XX  
PF 10-SEP-1997; 97US-0927230.  
XX  
PR 10-SEP-1997; 97US-0927230.  
XX  
PA (UYFL) UNIV FLORIDA.  
XX  
PI Silverman DN, Nick HS;  
XX  
XX WPI; 2000-578537/54.  
XX  
DR P-PSDB; AAB08204.  
XX  
DR  
XX  
XX Novel human manganese superoxide dismutase protein useful as  
XX antioxidant for treating cytotoxicity caused by superoxide radicals,  
XX inflammation and reperfusion injury following ischemia -  
XX  
XX Disclosure; Columns 25-26; 27pp; English.  
XX  
XX The present sequence encodes a human manganese superoxide dismutase  
XX protein (hMn SOD). The specification describes a modified hMn SOD which  
XX has a catalytic activity which differs from natural hMn SOD in that it  
XX exhibits reduced or no product inhibition compared to natural hMn  
XX SOD. The natural hMn SOD is especially modified with one or more  
XX substitutions in amino acids 26, 30, 34, 74, 143, 159, 161 and 163.  
XX The modified hMn SOD is useful for protecting a cell from damage caused  
XX by superoxide radicals and for treating a subject suffering from  
XX cytotoxicity caused by superoxide radicals. The protein is useful for  
XX treating inflammation and reperfusion injury following ischemia.  
XX  
SQ Sequence 813 BP; 218 A; 203 C; 214 G; 178 T; 0 other;  
Alignment Scores:  
Pred. No.: 0.000928 Length: 813  
Score: 74.00 Matches: 12  
Percent Similarity: 78.57% Conservative: 10  
Best Local Similarity: 42.86% Mismatches: 6  
Query Match: 51.03% Indels: 0  
DB: 21 Gaps: 0  
US-09-987-190-2 (1-30) x AAA63891 (1-813)  
QY 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20  
Db 115 AAGCAGACGCTCCCGACCTGCGCTACGACGCGCCCTGGAACCTCAACGCG 174  
QY 21 GlnIleAsnGluIle\*\*\*TyrThr 28  
Db 175 CAGATCATGCAGCTGCACACGAC 198  
RESULT 31  
ABA94453  
ID ABA94453 standard; cDNA; 849 BP.  
XX  
XX ABA94453;  
09-APR-2002 (first entry)  
XX  
DE Human manganese superoxide dismutase (hSODm) encoding cDNA.  
XX  
KW Human manganese superoxide dismutase; antiinflammatory; antiarthritic;  
KW antirheumatic; vasotropic; osteopathic; hSODm; human; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 100..768  
FT /tag= a  
FT /product= "hSODm"  
FT 100..171  
FT sig\_peptide 100..171  
FT /tag= b  
FT mat\_peptide 172..765  
FT

```

FT XX /+tag= c
XX PN US6326003-B1.
XX PD 04-DEC-2001.
XX PF 14-OCT-1986; 86US-0918534.
XX PR 14-OCT-1986; 86US-0918534.
XX PA (CHIR ) CHIRON CORP.
XX PI Halliwell RA, Bell GT, Mullenbach GT;
XX DR WPI; 2002-129495/17.
XX DR P-PSDB; ABB07330.
XX PT Treating inflammatory joint disease such as osteoarthritis, rheumatoid
XX PT arthritis and post ischemic tissue injury, comprises injecting human
XX XX manganese superoxide dismutase -
XX XX
XX XX Example 2; Fig 4; 14pp; English.
XX XX
XX CC The invention relates to a method of treating a patient with inflammatory
XX CC joint disease. The method involves intra-articularly injecting into the
XX CC afflicted joint a solution of human manganese superoxide dismutase
XX CC (hSODm). The treatment is used to treat an inflammatory joint disease,
XX CC particularly osteoarthritis or rheumatoid arthritis. The method can also
XX CC be used to treat post ischemic tissue injury. The present sequence
XX CC represents a cDNA encoding the hSODm.
XX XX
XX SQ Sequence 849 BP; 256 A; 197 C; 227 G; 169 T; 0 other;

Alignment Scores:
Pred. No.: 0.00098 Length: 849
Score: 74.00 Matches: 12
Percent Similarity: 78.57% Conservative: 10
Best Local Similarity: 42.86% Mismatches: 6
Query Match: 51.03% Indels: 0
DB: 24 Gaps: 0

US-09-987-190-2 (1-30) x ABA94453 (1-849)
QY 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20
Db 172 AAGCAGACGCTTGGCCGAGCTTGCCTGACGACGCGCTGGAACCTCAGATCAGCCG 231
21 GlnIleAsnGluIle**TyrThr 28
232 CAGATCATGCACTGCACACACG 255

RESULT 32
AAQ94279
ID AAQ94279 standard; cDNA; 966 BP.
XX AC AAQ94279;
XX XX
XX XX 23-MAY-1996 (first entry)
XX DE Human manganese superoxide dismutase (Gln29) CDS and partial 3'-UTR.
XX KW hMn-SOD; manganese; superoxide dismutase; tetramer; inflammation;
XX KM autoimmune disease; rheumatoid arthritis; neoplasm; emphysema;
XX XX osteoarthritis; wound healing; ss.
XX OS Homo sapiens.
XX XX
XX XX Key Location/Qualifiers
XX FH mat_peptide 1..594
XX FT /+tag= a
XX FT /product= hMn-SOD (Gln29)
XX PN EP676472-A1.

```

```

XX XX 11-OCT-1995.
XX PD 10-MAR-1988; 88EP-0107460.
XX PF 24-DEC-1987; 87DE-3744038.
XX PR 14-MAR-1987; 87DE-3708306.
XX PR 26-MAY-1987; 87DE-3717695.
XX PR 10-JUL-1987; 87DE-3722884.
XX PA (BOEH ) BOEHRINGER INGELHEIM INT GMBH.
XX PI Hauptmann R, Heckl K, Krystek E, Maurer-Fogy I;
XX PI Ostermann E, Spevak W, Stralow C, Wiche-Castanon M;
XX PI Zoepfel A;
XX DR WPI; 1995-346092/45.
XX XX
XX XX Genes encoding recombinant human manganese superoxide dismutase -
XX PT for treatment, prevention and diagnosis of inflammatory diseases
XX XX
XX PS Claim 8; Page 6 and Page 18; 54pp; German.
XX XX
XX CC The present sequence codes for the 198 amino acid long hMn-SOD
XX CC mature polypeptide having a Gln residue at position 29. The cDNA
XX CC can be inserted into expression vectors containing appropriate
XX CC signal sequences, etc. for expression of recombinant hMn-SOD.
XX CC Mature hMn-SOD is useful for treating, preventing or diagnosing
XX CC inflammatory, degenerative, neoplastic and rheumatic disorders;
XX CC in wound healing and in autoimmune diseases.
XX XX
XX SQ Sequence 966 BP; 279 A; 191 C; 216 G; 280 T; 0 other;

Alignment Scores:
Pred. No.: 0.00116 Length: 966
Score: 74.00 Matches: 12
Percent Similarity: 78.57% Conservative: 10
Best Local Similarity: 42.86% Mismatches: 6
Query Match: 51.03% Indels: 0
DB: 16 Gaps: 0

US-09-987-190-2 (1-30) x AAQ94279 (1-966)
QY 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20
Db 1 AAGCAGCTTTGGCCGAGCTTGCCTGACGACGCGCTGGAACCAACATCAATGCT 60
21 GlnIleAsnGluIle**TyrThr 28
61 CAATCATGCAATGCACCACTCT 84

RESULT 33
AAT31018
ID AAT31018 standard; cDNA to mRNA; 969 BP.
XX AC AAT31018;
XX XX
XX XX 26-SEP-1996 (first entry)
XX DE Human manganese superoxide dismutase cDNA.
XX KW Gene therapy; hypoxia related enhancer element; HREB; ischemia;
XX KM reperfusion; promoter; superoxide dismutase; ss.
XX XX
XX OS Homo sapiens.
XX XX
XX XX Key Location/Qualifiers
XX FH CDS 61..729
XX FT /+tag= a
XX FT W09620276-A1.
XX PN 04-JUL-1996.

```









PT treating inflammation or bronchial pulmonary dysplasia, reducing  
 PT reperfusion injury or prolonging organ survival  
 XX  
 PS Example 8; Fig 6A-6B; 27pp; English.  
 XX  
 CC DNA fragments (AA134278-82) cover the 6 exons of the human manganese  
 CC superoxide dismutase gene. The fragment given in AA134278 covers  
 CC the promoter region and exons 1 and 2. The MnSOD gene was isolated  
 CC from human placental DNA using a MnSOD cDNA clone (see also AA134277)  
 CC as probe. The isolated gene can be introduced into mammalian cell  
 CC hosts for the prodn. of MnSOD (see also AA000018), useful for the  
 CC treatment of e.g. ischaemia and inflammation.  
 XX  
 SQ Sequence 1007 BP; 154 A; 330 C; 397 G; 126 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 0.00122 Length: 1007  
 Score: 74.00 Matches: 12  
 Percent Similarity: 78.57% Conservative: 10  
 Best Local Similarity: 42.86% Mismatches: 6  
 Query Match: 51.03% Indels: 0  
 DB: 17 Gaps: 0  
 US-09-987-190-2 (1-30) x AA134278 (1-1007)  
 QY 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20  
 DB 833 AAGCAGCAGCCCTCCGACCTCCCTAGACTACGCGCCCTCGAACCACATCAACCGC 892  
 QY 21 GlnIleAsnGluIle\*\*\*TyrThr 28  
 DB 893 CAGATCATGCGACTGACACACAGC 916  
 RESULT 39  
 AA063914  
 ID AA063914 standard; DNA; 1008 BP.  
 AC AA063914;  
 XX  
 XX 21-JUN-1994 (first entry)  
 DT  
 XX  
 DE MnSOD DNA (exons 1 and 2).  
 XX  
 KM MnSOD; manganese superoxide dismutase; N-terminal; catalyst;  
 KM reperfusion injury following ischaemia; superoxide; SO; molecular oxygen;  
 KM anti-inflammatory; bs.  
 XX  
 XX Homo sapiens.  
 Key Location/Qualifiers  
 FT exon 421..510  
 FT /\*tag= a  
 FT 783..986  
 FT /\*tag= b  
 FT 74..102  
 FT stem\_loop  
 FT /\*tag= c  
 FT 214..233  
 FT /\*tag= d  
 FT 289..300  
 FT repeat\_unit  
 FT /\*tag= e  
 FT /rpt\_type= direct  
 FT /note= "repeat 3"  
 FT 405..415  
 FT /\*tag= f  
 FT /rpt\_type= direct  
 FT /note= "repeat 3"  
 FT 361..370  
 FT repeat\_unit  
 FT /\*tag= g  
 FT /rpt\_type= direct  
 FT /note= "repeat 4"  
 FT 373..383  
 FT repeat\_unit  
 FT /\*tag= h  
 FT /rpt\_type= direct

FT GC\_signal  
 FT /note= "repeat 4"  
 FT 71..76  
 FT /\*tag= i  
 FT /note= "SP1 binding site"  
 FT 247..257  
 FT /\*tag= j  
 FT /note= "SP1 binding site"  
 FT 296..307  
 FT /\*tag= k  
 FT /note= "SP1 binding site"  
 FT 324..334  
 FT /\*tag= l  
 FT /note= "SP1 binding site"  
 FT 1008  
 FT misc\_RNA  
 FT /\*tag= m  
 FT /note= "Represents sequence of undefined length  
 FT between exons 2 and 3."  
 FT US5270195-A.  
 FT 14-DEC-1993.  
 FT 22-NOV-1985; 85US-0801090.  
 FT 29-OCT-1986; 86IE-0002851.  
 FT (BIOT-) BIO-TECHNOLOGY GENERAL CORP.  
 FT Beck Y, Hartman JR;  
 FT WPI; 1993-404931/50.  
 FT Expression plasmid in Escherichia coli host system - encodes  
 FT human manganese superoxidisedismutase analogue, useful for e.g.  
 FT treating inflammation  
 FT Disclosure; Fig 6a-6d; 27pp; English.  
 XX  
 XX The sequence shows part of the genomic DNA that encodes a  
 XX manganese superoxide dismutase which can be used  
 XX to catalyse the reduction of superoxide (SO) radicals to hydrogen  
 XX peroxide and molecular oxygen. It can be used to reduce  
 XX reperfusion injury following ischaemia and prolong the survival of  
 XX excised organs. It can also be used as a long acting  
 XX anti-inflammatory drug. See sequences (AA063914-17) for other fragments.  
 XX  
 SQ Sequence 1008 BP; 154 A; 330 C; 397 G; 126 T; 1 other;  
 Alignment Scores:  
 Pred. No.: 0.00122 Length: 1008  
 Score: 74.00 Matches: 12  
 Percent Similarity: 78.57% Conservative: 10  
 Best Local Similarity: 42.86% Mismatches: 6  
 Query Match: 51.03% Indels: 0  
 DB: 14 Gaps: 0  
 US-09-987-190-2 (1-30) x AA063914 (1-1008)  
 QY 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20  
 DB 833 AAGCAGCAGCCCTCCGACCTCCCTAGACTACGCGCCCTCGAACCACATCAACCGC 892  
 QY 21 GlnIleAsnGluIle\*\*\*TyrThr 28  
 DB 893 CAGATCATGCGACTGACACACAGC 916  
 RESULT 40  
 ABK84524  
 ID ABK84524 standard; cDNA; 1026 BP.  
 AC ABK84524;  
 DT 14-AUG-2002 (first entry)

XX Human cDNA differentially expressed in granulocytic cells #1095.  
 DE  
 XX  
 KW Human; se; granulocytic cell; DNA chip; bacterial infection;  
 KW viral infection; parasitic infection; protozoal infection;  
 KW fungal infection; sterile inflammatory disease; psoriasis;  
 KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;  
 KW cardiac reperfusion injury; renal reperfusion injury; ARDS;  
 KW adult respiratory distress syndrome; inflammatory bowel disease;  
 KW Crohn's disease; ulcerative colitis; periodontal disease;  
 KW granulocyte activation; chronic inflammation; allergy.  
 KW  
 XX Homo sapiens.  
 OS  
 XX WO200228999-A2.  
 PN  
 XX 11-APR-2002.  
 PD  
 XX 03-OCT-2001; 2001WO-US30821.  
 PF  
 XX 03-OCT-2000; 2000US-237189P.  
 PA (GENE-) GENE LOGIC INC.  
 PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;  
 XX WPI; 2002-435328/46.  
 DR  
 XX  
 DR  
 XX  
 PT Detecting granulocyte activation by detecting differential expression  
 PT of genes associated with granulocyte activation, which serves as  
 PT diagnostic markers that is useful for monitoring disease states and  
 PT drug toxicity -  
 XX  
 PS Claim 1; SEQ ID No 1095; 114bp; English.  
 XX  
 CC The invention relates to detecting (M1) granulocyte (GC) activation  
 CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by  
 CC DNA chip analysis as given in the specification, and comparing  
 CC the expression level to an expression level in an unactivated  
 CC GC, where differential expression of Gs is indicative of GCA.  
 CC Also included are modulating (M2) GA by contacting GC with an agent  
 CC that alters the expression of at least one gene in Gs; (2) screening (M3)  
 CC for an agent capable of modulating GCA or an inflammation (especially  
 CC chronic) in a tissue, an allergic response in a subject, exposure of a  
 CC subject to a pathogen or sterile inflammatory disease using the  
 CC gene expression profile; (3) detecting (M4) an inflammation (especially  
 CC chronic) in a tissue, an allergic response in a subject, exposure of a  
 CC subject to a pathogen or sterile inflammatory disease, by detecting the  
 CC level of expression in a sample of the tissue of gene(s) from Gs, where  
 CC the level of expression of the gene is indicative of inflammation;  
 CC (4) treating (M5) an inflammation (especially chronic) or in a tissue,  
 CC an allergic response in a subject, exposure of a subject to a pathogen  
 CC or sterile inflammatory disease, by contacting a tissue having  
 CC inflammation with an agent that modulates the expression of gene(s)  
 CC from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for  
 CC modulating GA; M3 is useful for screening an agent capable of modulating  
 CC GCA preferably in an inflammation in a tissue; M4 is useful for  
 CC detecting an inflammation (especially chronic) in a tissue, an allergic  
 CC response in a subject, exposure of a subject to a pathogen or sterile  
 CC inflammatory disease (e.g. psoriasis, rheumatoid arthritis,  
 CC glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal  
 CC reperfusion injury, ARDS, adult respiratory distress syndrome,  
 CC inflammatory bowel disease, Crohn's disease, ulcerative colitis,  
 CC periodontal disease; also bacterial infection, viral infection,  
 CC parasitic infection, protozoal infection, fungal infection and M5 is  
 CC useful for treating one of the above conditions. The present  
 CC sequence represents a gene differentially expressed in granulocytes.  
 CC Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 CC Sequence 1026 BP; 283 A; 216 C; 245 G; 282 T; 0 other;  
 CC  
 XX

Alignment Scores:  
 Pred. No.: 0.00125 Length: 1026  
 Score: 74.00 Matches: 12  
 Percent Similarity: 78.57% Conservative: 10  
 Best Local Similarity: 42.86% Mismatches: 6  
 Query Match: 51.03% Indels: 0  
 DB: 24 Gaps: 0  
 US-09-987-190-2 (1-30) x ABK84524 (1-1026)  
 QY 1 LysTyrSerLeuPProGluLeuAspTyrGluPheSerAlaThrGluProTyrTlleSerGly 20  
 Db 77 AAGCAGAGCCTCCCGGACCTGCGCTTACGACTACGCGCCCTCGAACTCACATCACGCG 136  
 QY 21 GlnIleAsnGluIle\*\*\*TyrThr 28  
 Db 137 CAGATCATGCAGCTGCACACGACG 160

Search completed: April 8, 2003, 02:21:59  
 Job time : 65.422 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 9, 2003, 14:01:45 ; Search time 3.99083 Seconds  
(without alignments)  
311.788 Million cell updates/sec

Title: US-09-987-190-2  
Perfect score: 30  
Sequence: 1 KYSLPELDYFSAATEPYISQINEIXYTX 30

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Indexed: 112892 seqs, 41476328 residues

Word size: 0

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database: SwissProt.40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Length	ID	Description
1	25	83.3	234 1	SODM_CANAL
2	10	33.3	233 1	SODM_YEAST
3	8	26.7	206 1	SODM_MYCLP
4	8	26.7	206 1	SODM_MYCLP
5	7	23.3	206 1	SODM_MYCFO
6	7	23.3	206 1	SODM_MYCSM
7	7	23.3	206 1	SODM_MYCSC
8	7	23.3	207 1	SODM_MYCTU
9	7	23.3	207 1	SODM_MYCTU
10	7	23.3	488 1	GCS2_BACSV
11	6	23.0	1034 1	CAP2_SYNY3
12	6	20.0	137 1	Y142_LEUWC
13	6	20.0	137 1	UK14_HUMAN
14	6	20.0	201 1	SODM_PROFR
15	6	20.0	207 1	SMGH_LISMO
16	6	20.0	207 1	SODM_CHLBN
17	6	20.0	278 1	PURU_HAELN
18	6	20.0	334 1	MENC_MYCLE
19	6	20.0	347 1	UTR2_YEAST
20	6	20.0	360 1	ID12_HALNI
21	6	20.0	428 1	HEMY_HAELN
22	6	20.0	535 1	TCE1_AVEBA
23	6	20.0	535 1	TCE2_AVEBA
24	6	20.0	561 1	YDDA_ECOLI
25	6	20.0	879 1	BALI_MOUSE
26	6	20.0	881 1	BALI_MOUSE
27	6	20.0	948 1	UVR4_HUMAN
28	6	20.0	949 1	UVR4_NEIMA
29	6	20.0	950 1	UVR4_NEIMA
30	6	20.0	1325 1	BCC3_ACEXY
31	5	16.7	74 1	VI94_FOWPV
32	5	16.7	85 1	YKK2_CAEEL
33	5	16.7	101 1	YCK5_CHLEA
	5	16.7	109 1	Y265_UREPA

34	5	16.7	115 1	ANFC_SCYCA	P23259 scyliorhinu
35	5	16.7	115 1	SMD2_SCHPO	014036 schizosacch
36	5	16.7	124 1	REV_SIVCZ	P17280 chimpanzee
37	5	16.7	125 1	Y364_AQUAE	066689 aquifex aeo
38	5	16.7	127 1	Y026_NPYOP	010297 oryza pseu
39	5	16.7	129 1	VAL2_ABMVM	P21944 abutilon mo
40	5	16.7	129 1	VAL2_TMOV	006658 tomato mott
41	5	16.7	132 1	VAL3_PHVU	006925 pepper huas
42	5	16.7	133 1	TATB_VIBCH	P57063 vibrio chol
43	5	16.7	137 1	UK14_CAPHI	P80601 capra hircu
44	5	16.7	138 1	EXD1_VIBCH	052044 vibrio chol
45	5	16.7	140 1	LCA_TIRIVU	029145 trichosurus
46	5	16.7	141 1	APGB_MOUSE	09C9Y1 mus musculu
47	5	16.7	144 1	VE6_HPV54	Q81018 human papil
48	5	16.7	145 1	YB6_YEAST	P38210 saccharomyc
49	5	16.7	149 1	ARGR_BACBD	Q9K973 bacillus ha
50	5	16.7	149 1	ARGR_BACLI	086130 bacillus li
51	5	16.7	149 1	ARGR_BACST	031408 bacillus st
52	5	16.7	149 1	ARGR_BACSV	P17893 bacillus su
53	5	16.7	157 1	IBP_BUCAI	P57640 buchnera ap
54	5	16.7	162 1	YLF6_CAEEL	Q03598 caenorhabdi
55	5	16.7	165 1	RIMM_RICCN	Q32167 rickettsia
56	5	16.7	169 1	CALB_HUMAN	P06705 homo sapien
57	5	16.7	169 1	CALB_MOUSE	Q63810 mus musculu
58	5	16.7	170 1	CALB_DROME	P48451 drosophila
59	5	16.7	170 1	CALC_DROME	Q24214 drosophila
60	5	16.7	171 1	Y162_MERTH	Q27890 methanobact
61	5	16.7	173 1	HR12_MOUSE	P22760 mus musculu
62	5	16.7	175 1	CALC_RAT	P28470 rattus norv
63	5	16.7	175 1	FLAW_KLEPN	P04668 klebsiella
64	5	16.7	175 1	NU6W_FELCA	P48928 felis silve
65	5	16.7	177 1	RELX_MESAU	Q64171 mesocricetu
66	5	16.7	177 1	YDAB_MYCBO	P66830 mycobacteri
67	5	16.7	180 1	TAAS_MHEAT	P16347 triticum ae
68	5	16.7	180 1	Y788_RICPR	005978 rickettsia
69	5	16.7	182 1	RIMM_VIBCH	Q9K973 vibrio chol
70	5	16.7	185 1	RELI_HUMAN	P04808 homo sapien
71	5	16.7	185 1	RELX_MOUSE	P47932 mus musculu
72	5	16.7	188 1	V230_FOWPV	Q93505 fowlpox vir
73	5	16.7	191 1	KIP1_HUMAN	Q99828 homo sapien
74	5	16.7	191 1	KIP1_MOUSE	Q920F4 mus musculu
75	5	16.7	191 1	KIP1_RAT	Q9T010 rattus norv
76	5	16.7	200 1	NUPL_XENLA	P05221 xenopus lae
77	5	16.7	200 1	SODM_AGABI	Q9P4C6 agaricus bi
78	5	16.7	202 1	SODF_MERTH	Q60056 methanobact
79	5	16.7	203 1	TAAS_HORVU	P07596 hordium vul
80	5	16.7	203 1	PTH_ANASP	Q8Y744 anabena sp
81	5	16.7	207 1	YNO2_CAEEL	Q41965 caenorhabdi
82	5	16.7	208 1	KGUA_HAELN	P44310 haemophilus
83	5	16.7	210 1	SODF_SUDAC	Q08713 sulfolobus
84	5	16.7	210 1	SODF_SUDSO	P08857 sulfolobus
85	5	16.7	211 1	FIXI_AZOCA	Q66817 azotirizobiu
86	5	16.7	211 1	SODF_ACTAM	Q9P913 acidianus a
87	5	16.7	211 1	SODF_PYRAE	Q93724 pyrobaculum
88	5	16.7	212 1	UL03_HAVER	P28942 equine help
89	5	16.7	214 1	GTH3_ARATH	P42761 arabidopsis
90	5	16.7	216 1	LUXM_VIRBA	P54298 vibrio harv
91	5	16.7	225 1	TN12_MOUSE	Q54907 mus musculu
92	5	16.7	226 1	DTXR_CORDI	P33120 corynebacte
93	5	16.7	226 1	NHB2_RHORI	P29379 rhodococcus
94	5	16.7	228 1	RPE_STROO	Q41025 streptococc
95	5	16.7	231 1	FLGD_ECOLI	P75936 escherichia
96	5	16.7	232 1	FLGD_SALTY	P16321 salmonella
97	5	16.7	236 1	YREP_ECOLI	P76359 escherichia
98	5	16.7	239 1	DCUR_ECOLI	P39271 escherichia
99	5	16.7	242 1	TPIS_MYCFL	P48779 mycoplasma
100	5	16.7	243 1	DP3E_ECOLI	P30307 escherichia
101	5	16.7	243 1	YC22_MERJA	Q68641 methanococc
102	5	16.7	244 1	T2FC_YEAST	P55189 saccharomyc
103	5	16.7	244 1	Y181_MYCTU	Q07428 mycobacteri
104	5	16.7	245 1	RNPH_BACSV	P28619 bacillus su
105	5	16.7	245 1	SODM_NEUCR	Q9Y788 neurospora
106	5	16.7	247 1	TRAF_ECOLI	P14497 escherichia

107	5	16.7	249	1	TN12_HUMAN	043508	homo sapien	180	5	16.7	362	1	AROB_HAEN	P43879	haemophilus
108	5	16.7	250	1	TRMD_TREPA	083878	treponema p	181	5	16.7	363	1	MBHS_BRJA	P12635	bradyrhizob
109	5	16.7	251	1	BIOC_ECOLI	P12999	escherichia	182	5	16.7	366	1	YN10_YEAS	P33938	saccharomyc
110	5	16.7	253	1	BIOC_ERWHE	066898	erwinia her	183	5	16.7	372	1	PRS2_MERTM	P42811	methanobact
111	5	16.7	253	1	DF10_YEAST	P40526	saccharomyc	184	5	16.7	374	1	RECA_STRCO	050487	streptomyce
112	5	16.7	259	1	Y500_METUA	Q57923	methanococc	185	5	16.7	374	1	RECA_STRLI	P48294	streptomyce
113	5	16.7	260	1	RCEH_RHOSH	P11846	rhodobacter	186	5	16.7	374	1	RFAG_ECOLI	P25740	escherichia
114	5	16.7	260	1	S3AD_STAM	P04827	staphylococc	187	5	16.7	375	1	MBHS_CITFR	Q46045	citrobacter
115	5	16.7	260	1	VBRL_ICMV	008595	indian caas	188	5	16.7	375	1	PROB_SYNY3	P73071	synecocyst
116	5	16.7	269	1	AAKB_RAT	P60386	rattus norv	189	5	16.7	375	1	SOX3_MOUSE	P53782	mus musculu
117	5	16.7	272	1	Y439_MYCGE	P47677	mycoplasma	190	5	16.7	376	1	HOSC_THETH	087718	thermus the
118	5	16.7	276	1	YJIC_ECOLI	P33374	escherichia	191	5	16.7	376	1	PEXE_MOUSE	Q97080	mus musculu
119	5	16.7	281	1	SODP_BACSU	035023	bacillus su	192	5	16.7	377	1	MYG_NORMA	P31537	nordictus ma
120	5	16.7	288	1	HEMK_MYCLE	P45832	mycobacteri	193	5	16.7	377	1	PEXE_HUMAN	Q75381	homo sapien
121	5	16.7	289	1	ATRG_PASWU	P916b6	pasteurella	194	5	16.7	379	1	WZA_ECOLI	P76388	escherichia
122	5	16.7	298	1	CD38_RABIT	Q9m203	oryctolagus	195	5	16.7	379	1	WZA_SALTI	Q82554	salmonella
123	5	16.7	299	1	RR42_SCHPO	060124	schizosacch	196	5	16.7	379	1	WZA_SALTY	Q82nq9	salmonella
124	5	16.7	302	1	Y091_METUA	Q57556	methanococc	197	5	16.7	387	1	Y4PF_RHISN	P35615	rhizobium s
125	5	16.7	304	1	HEMK_MYCTU	Q10602	mycobacteri	198	5	16.7	388	1	FETB_MOUSE	Q9qxc1	mus musculu
126	5	16.7	304	1	HEMZ_THEVO	Q97809	thermoplas	199	5	16.7	389	1	MTSA_LACTC	P48717	lactococcus
127	5	16.7	304	1	NPX4_RAT	Q932n4	rattus norv	200	5	16.7	390	1	TRPB_MERTM	P26921	methanobact
128	5	16.7	305	1	PIR_ECOLI	P03067	escherichia	201	5	16.7	392	1	PSMR_SUITO	Q975u2	sulfobolus
129	5	16.7	308	1	HEMZ_THENC	Q911b8	thermoplas	202	5	16.7	394	1	DUS4_HUMAN	Q31115	homo sapien
130	5	16.7	309	1	SOX3_XENLA	P55863	xenopus lae	203	5	16.7	396	1	APA4_HUMAN	P53_ONCMY	oncorhynch
131	5	16.7	310	1	PYRB_HALNL	Q9phn4	halobacteri	204	5	16.7	396	1	P53_ONCMY	P25035	oncorhynch
132	5	16.7	310	1	PYRB_LACTA	Q9cft79	lactococcus	205	5	16.7	398	1	A23D_DROME	Q44093	drosophila
133	5	16.7	310	1	VCAP_BPT3	P20324	bacterioph	206	5	16.7	398	1	FTS2_WOLSP	P54855	wolbachia s
134	5	16.7	311	1	MRAP_BACSU	Q07876	bacillus su	207	5	16.7	409	1	YOPM_YERPE	P17778	yersinia pe
135	5	16.7	312	1	SOX3_CHICK	P48433	gallus gall	208	5	16.7	411	1	EP1G_TRYCR	P34715	trypanosoma
136	5	16.7	312	1	YDIR_ECOLI	P77378	escherichia	209	5	16.7	411	1	LCYB_SYNP7	Q55276	synecococc
137	5	16.7	313	1	FIXB_ECOLI	P15074	escherichia	210	5	16.7	412	1	PGKP_PLAPA	P02893	plasmodium
138	5	16.7	315	1	PHSS_DESBA	P13063	desulfovibr	211	5	16.7	412	1	PGKP_ALCEU	P03020	alcaligenes
139	5	16.7	316	1	BGM1_LACAC	Q07685	lactobacilli	212	5	16.7	412	1	RBAL_CAREL	P50319	caenorhabdi
140	5	16.7	316	1	CBT_ECOLI	Q47083	escherichia	213	5	16.7	413	1	PGKC_ALCEU	P50319	alcaligenes
141	5	16.7	318	1	ANX4_CANRA	P50994	canis famli	214	5	16.7	415	1	YO12_YEAST	P35193	saccharomyc
142	5	16.7	319	1	YQOP_BACSU	P54553	bacillus su	215	5	16.7	416	1	TR16_CHICK	P18519	gallus gall
143	5	16.7	320	1	BIRA_SALTY	P37416	salmonella	216	5	16.7	417	1	SAHH_SULSO	P50252	salifolobus
144	5	16.7	321	1	BIRA_ECOLI	P06709	escherichia	217	5	16.7	417	1	TR16_MOUSE	Q920u1	mus musculu
145	5	16.7	322	1	YVBF_ECOLI	P77757	escherichia	218	5	16.7	418	1	EP1G_ORYSA	Q92r17	oryza sativ
146	5	16.7	324	1	GLXA_RHIME	Q87389	rhizobium m	219	5	16.7	419	1	GLPB_ECOLI	P13033	escherichia
147	5	16.7	325	1	YRBG_ECOLI	P43394	escherichia	220	5	16.7	423	1	TIG_BACSU	P80658	bacillus su
148	5	16.7	326	1	MENC_MYCTU	Q06419	mycobacteri	221	5	16.7	424	1	BM10_HUMAN	Q95353	homo sapien
149	5	16.7	330	1	PG11_ARATH	Q9m5j9	arabidopsis	222	5	16.7	424	1	CSP_PLAPT	P13814	plasmodium
150	5	16.7	330	1	PG12_ARATH	Q9m5j8	arabidopsis	223	5	16.7	425	1	K2C1_XENLA	P42611	rhodobacter
151	5	16.7	331	1	GALR_STRTR	Q9yb11	streptococc	224	5	16.7	425	1	TR16_RAT	Q9k8f3	bacillus ha
152	5	16.7	333	1	PAP2_VACCC	P21033	vaccinia vi	225	5	16.7	427	1	TIG_LACTA	P48077	cyanothora
153	5	16.7	333	1	PAP2_VACCV	P07617	vaccinia vi	226	5	16.7	427	1	TR16_HUMAN	P54444	bacillus su
154	5	16.7	333	1	PAP2_VARY	P33052	variola vir	227	5	16.7	428	1	APG2_METUA	P35158	saccharomyc
155	5	16.7	337	1	AL13_SCHPO	Q13953	schizosacch	228	5	16.7	428	1	BCH2_RHOCA	Q86171	rhodobacter
156	5	16.7	338	1	GPDA_STRPN	Q97n61	streptococc	229	5	16.7	428	1	HEM1_ANASP	O08333	anabaena sp
157	5	16.7	343	1	LICH_ACICA	Q43961	actinobact	230	5	16.7	431	1	TIG_BACHD	Q9k8f3	bacillus ha
158	5	16.7	344	1	MBHS_AZCOH	P11990	azotobacter	231	5	16.7	432	1	HEM1_CYPAP	P48077	cyanothora
159	5	16.7	346	1	ADH_MYCTU	P11975	mycobacteri	232	5	16.7	432	1	YRQO_BACSU	P54444	bacillus su
160	5	16.7	347	1	ID12_LACTA	Q9c1f5	lactococcus	233	5	16.7	433	1	MTM1_YEAST	P35158	saccharomyc
161	5	16.7	348	1	NTRB_PROVU	P28788	proteus vul	234	5	16.7	435	1	HEM1_LISMO	O8y6x4	listeria mo
162	5	16.7	349	1	NTRB_ECOLI	P06712	escherichia	235	5	16.7	437	1	HEM1_PAPMA	Q69190	paenibacilli
163	5	16.7	349	1	NTRB_KLEPN	P06218	klebsiella	236	5	16.7	439	1	NMT1_PSEAE	Q97r33	pseudomonas
164	5	16.7	349	1	NTRB_GALTY	P41788	salmonella	237	5	16.7	439	1	YD75_MYCTU	P11803	mycobacteri
165	5	16.7	349	1	US30_HCMYA	P09706	human cytom	238	5	16.7	443	1	SOX3_HUMAN	P41225	homo sapien
166	5	16.7	350	1	POT2_HAENI	P44731	haemophilus	239	5	16.7	444	1	C120_SYNY3	Q59990	synecocyst
167	5	16.7	351	1	Y098_SYNY3	Q58860	synecocyst	240	5	16.7	445	1	DCDA_CORGL	Q02220	saccharomyc
168	5	16.7	351	1	CBIG_SALTY	Q05631	salmonella	241	5	16.7	447	1	YKYS_YEAST	P29955	xanthomonas
169	5	16.7	351	1	ID12_RHITO	Q98915	rhizobium 1	242	5	16.7	448	1	XANA_XANCP	P29955	xanthomonas
170	5	16.7	352	1	EGSA_THENC	Q9h1j6	thermoplas	243	5	16.7	448	1	YB00_METUA	O585f0	methanococc
171	5	16.7	352	1	RECA_CLOPE	P94666	clostridium	244	5	16.7	451	1	PTKE_HOWAN	Q13882	homo sapien
172	5	16.7	352	1	VATC_DEIRA	Q97mho	deinococcus	245	5	16.7	452	1	HEMN_RHOSH	P33770	rhodobacter
173	5	16.7	353	1	DRN2_MOUSE	P56542	mus musculu	246	5	16.7	452	1	SELA_AQUAE	P95651	rhodobacter
174	5	16.7	356	1	AROB_CLOAB	Q97km3	clostridium	247	5	16.7	452	1	SELA_AQUAE	O67140	aquifex ao
175	5	16.7	356	1	SMRI_PODAN	O08142	podopora a	248	5	16.7	457	1	PVAL_MOUSE	P28650	mus musculu
176	5	16.7	358	1	MBHS_AZOVI	P21950	azotobacter	249	5	16.7	458	1	NU4M_STRCA	Q21466	struthio ca
177	5	16.7	360	1	MBHS_RHILV	P18637	rhizobium 1	250	5	16.7	459	1	NU4M_CHICK	P18939	gallus gall
178	5	16.7	360	1	POT1_HAENI	P41668	haemophilus	251	5	16.7	459	1	NU4M_DASNO	O21334	dasyus nov
179	5	16.7	360	1	VP3_ARMV	P24820	arabid msa	252	5	16.7	459	1	NU4M_MOUSE	P03911	mus musculu

253	5	16.7	459	1	NU4M_POIOR	Q95917	polypetrus	580	1	YML1_ARATH	O22752	arabidopsis
254	5	16.7	459	1	NU4M_RABIT	O79436	oryctolagus	581	1	KAPP_ARATH	P46014	arabidopsis
255	5	16.7	459	1	NU4M_RAT	P05508	rattus norv	583	1	FRDA_MYCTH	O10760	mycobacteri
256	5	16.7	460	1	NU4M_SCYCA	O79410	scylliorhinu	599	1	CENE_HUMAN	P07139	homo sapien
257	5	16.7	460	1	NU4M_SQUAC	O92445	squnulus aca	599	1	CENE_MOUSE	P27790	mus muscucu
258	5	16.7	461	1	NU4M_XENLA	P03912	xenopus lae	606	1	CENE_CRIGR	P48998	cricketulus
259	5	16.7	462	1	YMO1_YEAST	Q04461	saccharomyc	606	1	Y4BO_RHISN	P55302	thizobium s
260	5	16.7	463	1	SYS_AERPE	Q9yag3	aeropyrum p	621	1	FOR_THETI	O56303	thermococcu
261	5	16.7	464	1	IRLS_BURPS	O31396	burkholderi	625	1	BGAL_LACSK	O48846	lactobacilli
262	5	16.7	465	1	BGLC_RHIME	O92467	rhizobium m	626	1	GIDA_CLOAB	O97843	clostridium
263	5	16.7	466	1	SERA_SCHPO	P87228	schizosacch	627	1	GIGH_BACSU	P39118	bacillus su
264	5	16.7	468	1	LIP2_RAT	P54318	rattus norv	628	1	BGAL_LACAC	O07664	lactobacilli
265	5	16.7	469	1	GATB_THETH	Q91cx2	thermus the	628	1	LU_HUMAN	P50865	homo sapien
266	5	16.7	469	1	VL2_BPV1	P03109	bovine papl	629	1	GIDA_XYLFA	O9pbh4	xyliella fas
267	5	16.7	471	1	CA34_BOVIN	P28094	bos taurus	630	1	GIDA_CLOPE	O8xh31	clostridium
268	5	16.7	471	1	TIG_SYNY3	O55511	synechocyst	637	1	VATI_THEAC	O9mth8	thermoplasm
269	5	16.7	473	1	RUMA_RHOGE	O91p88	rhodocyclus	640	1	SP10_YEAST	P35208	saccharomyc
270	5	16.7	475	1	GUNA_CLOCE	P17901	clostridium	645	1	SP5D_BACSU	O03524	bacillus su
271	5	16.7	479	1	MP1P_DROME	P20483	drosophila	649	1	HEXB_STRPN	P14160	streptococc
272	5	16.7	480	1	GLTD_AZOBH	O05756	azospirillum	656	1	REP_BUCAP	O51889	buchnera ap
273	5	16.7	482	1	RNF9_HUMAN	Q9udv6	homo sapien	658	1	MUTL_LACLA	O99xm7	streptococc
274	5	16.7	483	1	GATB_RICPR	O9ze11	rickettsia	660	1	MUTL_STRPY	P53135	saccharomyc
275	5	16.7	484	1	SCR8_VIBAL	P13394	vibrio algi	668	1	YGLJ_YEAST	P57031	paenibacill
276	5	16.7	489	1	PSD3_DAUCA	Q06364	daucus caro	675	1	C1B9_PAPBP	P38738	saccharomyc
277	5	16.7	489	1	RNF9_MOUSE	Q9wuh5	mus muscucu	679	1	YHC9_YEAST	O94289	homo sapien
278	5	16.7	491	1	GLG2_SYNY3	P72623	synechocyst	681	1	GFA2_HUMAN	O92269	mus muscucu
279	5	16.7	491	1	ILVC_PASMU	O9clf1	pasteurella	681	1	GFA2_MOUSE	O11157	mycobacteri
280	5	16.7	492	1	SES3_HUMAN	P58005	homo sapien	684	1	Y492_MYCTU	P36447	chlamydomon
281	5	16.7	492	1	SES3_MOUSE	O9cyg7	mus muscucu	688	1	CHLE_CHIRE	P38448	saccharomyc
282	5	16.7	493	1	UGDH_MOUSE	O70475	mus muscucu	688	1	YB9F_YEAST	P38447	saccharomyc
283	5	16.7	493	1	UGDH_RAT	O70199	rattus norv	690	1	HEL5_MERTH	O26901	methanobact
284	5	16.7	494	1	UGDH_BOVIN	P12378	bos taurus	703	1	PEK_STRGR	O9ueh8	streptomyce
285	5	16.7	494	1	UGDH_HUMAN	O60701	homo sapien	705	1	DPO2_YEAST	P38121	saccharomyc
286	5	16.7	506	1	YC46_GUTTH	O78439	guillardia	707	1	HS88_NEUCR	O74225	neurospora
287	5	16.7	507	1	SXA2_SCHPO	P32825	schizosacch	712	1	DNLJ_RHOMR	P49421	rhodothermu
288	5	16.7	510	1	GEPD_ASPNG	P48826	aspergillus	715	1	HEL5_PYRAB	O9v089	pyrococcus
289	5	16.7	511	1	GEPD_EMENT	P41764	emeritella	721	1	HEL5_PYRUV	O73946	pyrococcus
290	5	16.7	515	1	CP12_HUMAN	P05177	homo sapien	721	1	RNR_UREPA	O9p718	ureaplasma
291	5	16.7	516	1	DHCR_HUMAN	O15392	homo sapien	728	1	YJ85_YEAST	P38121	saccharomyc
292	5	16.7	521	1	VGIC_HSVBC	P14378	bovine herp	731	1	TOPI_TREPA	O83409	creponema p
293	5	16.7	525	1	DIMH_CAREL	O17397	caenorhabdi	736	1	PPK_STRCO	O9kz46	streptomyce
294	5	16.7	528	1	CTK1_YEAST	Q03957	saccharomyc	746	1	RPR1_INCUJ	P19703	influenza c
295	5	16.7	528	1	PRPR_ECOLI	P77743	escherichia	756	1	EFER_HUMAN	O59014	homo sapien
296	5	16.7	531	1	CE27_HUMAN	Q02318	h cytochrom	761	1	TOPI_MERJA	O95164	methanococc
297	5	16.7	531	1	FXM1_SCHPO	O09752	schizosacch	765	1	YK07_YEAST	P36122	saccharomyc
298	5	16.7	532	1	IPAT_SHIFL	P18014	shigella fl	767	1	YNM4_YEAST	P53866	saccharomyc
299	5	16.7	535	1	CP27_RABIT	P17177	o cytochrom	768	1	KGP1_DROME	O03042	drosophila
300	5	16.7	537	1	CH61_MYCLE	P37578	mycobacteri	772	1	SYFB_UREPA	O9p943	ureaplasma
301	5	16.7	538	1	YAJ8_SCHPO	Q09908	schizosacch	774	1	NEK1_MOUSE	P51954	mus muscucu
302	5	16.7	541	1	PRPR_SALTY	P37702	arabidopsis	774	1	MEPV_HUMAN	O9hy58	pseudomonas
303	5	16.7	541	1	PRPR_SALTY	P74839	salmonella	781	1	YHY2_YEAST	O15553	homo sapien
304	5	16.7	543	1	RRP3_YEAST	P38712	saccharomyc	785	1	YHY2_YEAST	P38870	saccharomyc
305	5	16.7	544	1	MWR3_SINAT	P29092	sinapis alb	788	1	TRSI_HCMVA	O90665	human cytom
306	5	16.7	546	1	PCCB_SACER	P53003	saccharopol	790	1	LY14_YEAST	P40971	saccharomyc
307	5	16.7	546	1	TRCD_ARCHU	O28668	archaeoglob	791	1	KDGL_DROME	O01583	drosophila
308	5	16.7	548	1	CO9_MOUSE	P06683	mus muscucu	818	1	SYFB_RICCN	O92138	rickettsia
309	5	16.7	548	1	MYR0_BRANA	O00326	brassica na	819	1	YX07_CAREL	O11114	caenorhabdi
310	5	16.7	550	1	YMA0_YEAST	Q02322	saccharomyc	825	1	AKIH_SERMA	P27725	serratia ma
311	5	16.7	551	1	AGLA_RHIME	Q92318	rhizobium m	825	1	XEP_BIFRAN	O9aem9	bifidobacte
312	5	16.7	551	1	ILVD_PYRAB	Q9uz03	pyrococcus	830	1	RPOD_EUGGR	P23511	euglena gira
313	5	16.7	553	1	IAGA_SALTY	P43016	salmonella	831	1	NAH3_RAT	P26433	rattus norv
314	5	16.7	553	1	IAGA_SALTY	P43015	salmonella	832	1	NAH3_RABIT	P26432	oryctolagus
315	5	16.7	553	1	ILVD_MERJA	O58672	methanococc	833	1	HSF_YEAST	P10961	saccharomyc
316	5	16.7	558	1	SYFA_TREPA	O83938	treponema p	834	1	NAH1_HUMAN	P48754	homo sapien
317	5	16.7	558	1	LCB1_YEAST	P25045	saccharomyc	836	1	MCBL_SFVKA	P25950	shope fibro
318	5	16.7	561	1	CEP1_USITWA	P49602	ustilago ma	838	1	PAS_MOUSE	P19066	mus muscucu
319	5	16.7	562	1	EXG2_YEAST	P52911	saccharomyc	839	1	NAH3_DIDMA	O28383	didelphis m
320	5	16.7	562	1	SIS2_YEAST	P36024	saccharomyc	840	1	YHT1_YEAST	P38835	saccharomyc
321	5	16.7	563	1	LIP2_GEOON	P23294	geotrichum	841	1	IRG1_HCMVA	O42814	aspergillus
322	5	16.7	563	1	PRIM_THEMA	O9x1g3	thermotoga	846	1	ENV_HCMVA	P092715	human cytom
323	5	16.7	569	1	NODU_BRJJA	P26027	bradyrhizob	851	1	ENV_HVILB	P04562	human immun
324	5	16.7	570	1	IF37_CAREL	P30642	caenorhabdi	851	1	MUTS_STRPY	O99x18	streptococc
325	5	16.7	577	1	POF3_SCHPO	O74991	schizosacch	853	1	ENV_HVILMF	P19551	human immun

399	5	16.7	855	1	ENV_HV10Y	P20888 human immun	472	5	16.7	1400	1	RON_HUMAN	004912 homo sapien
400	5	16.7	856	1	ENV_HV1B1	P03375 human immun	473	5	16.7	1411	1	YM42_YEAST	003214 saccharomyc
401	5	16.7	856	1	ENV_HV1H2	P04578 human immun	474	5	16.7	1433	1	VGLM_BUNYM	P04505 bunywerera
402	5	16.7	856	1	ENV_HV1H3	P06624 human immun	475	5	16.7	1440	1	SYEP_HUMAN	P07781 homo sapien
403	5	16.7	856	1	ENV_HV1LM	P07626 human immun	476	5	16.7	1447	1	TOP2_DROME	P15384 drosophila
404	5	16.7	856	1	ENV_HV1LV	P03376 human immun	477	5	16.7	1474	1	A2MG_HUMAN	P10103 homo sapien
405	5	16.7	861	1	ENV_HV1BR	P03377 human immun	478	5	16.7	1482	1	P2P_HUMAN	P20742 homo sapien
406	5	16.7	861	1	SYL_HAE1R	P43827 haemophilus	479	5	16.7	1489	1	YGP0_YEAST	P53115 saccharomyc
407	5	16.7	861	1	ENV_HV1RH	P04579 human immun	480	5	16.7	1520	1	TOP2_CAEEL	Q23670 caenorhabdi
408	5	16.7	870	1	SYV_CAMDE	O99pe4 campylobact	481	5	16.7	1526	1	TP2A_CRIGR	P41515 cricetus
409	5	16.7	878	1	SYA_BUCAT	P57483 buchnera ap	482	5	16.7	1526	1	TP2A_RAT	P41516 rattus norv
410	5	16.7	892	1	TNPA_STRFR	P20189 streptomyc	483	5	16.7	1528	1	TP2A_MOUSE	Q01320 mus musculu
411	5	16.7	895	1	PMAI_CANAL	P28877 candida alb	484	5	16.7	1531	1	TP2A_HUMAN	P11388 homo sapien
412	5	16.7	906	1	LOXC_ARATH	P38418 arabidopsis	485	5	16.7	1533	1	TP2A_PIG	O46314 sus scrofa
413	5	16.7	900	1	MUTS_TREBA	O83348 treponema p	486	5	16.7	1547	1	TOP2_BOWMO	Q16140 dombuy mori
414	5	16.7	904	1	DPO1_MYCTU	O07700 mycobacteri	487	5	16.7	1670	1	CA34_HUMAN	Q01955 homo sapien
415	5	16.7	910	1	SYL_NEINA	O97w39 neisseria m	488	5	16.7	1695	1	KEP1_MOUSE	P33173 mus musculu
416	5	16.7	911	1	DPO1_MYCLE	P46835 mycobacteri	489	5	16.7	1758	1	CA24_CAEEL	P17140 caenorhabdi
417	5	16.7	911	1	ITH1_HUMAN	P19827 homo sapien	490	5	16.7	1763	1	CA24_ASCSU	P27353 ascatis buu
418	5	16.7	919	1	PWP2_HUMAN	O15269 homo sapien	491	5	16.7	1822	1	ITB4_HUMAN	P16144 homo sapien
419	5	16.7	925	1	GLHR_ANTEL	P35409 anthopleura	492	5	16.7	1829	1	RM1_MOUSE	O70472 mus musculu
420	5	16.7	928	1	KKRI_YEAST	P36003 saccharomyc	493	5	16.7	1829	1	Y296_HUMAN	O15015 homo sapien
421	5	16.7	935	1	ODOI_HAEIN	P45303 haemophilus	494	5	16.7	1888	1	YDT2_SCHPO	O14207 schizosacch
422	5	16.7	936	1	CDAS_HUMAN	O957h7 homo sapien	495	5	16.7	1906	1	KMLS_CHICK	P11799 gallus galli
423	5	16.7	936	1	MSH4_HUMAN	O15457 homo sapien	496	5	16.7	1916	1	R1F1_YEAST	P29539 saccharomyc
424	5	16.7	951	1	SYV_ECOLI	P07118 escherichia	497	5	16.7	1980	1	Y9B_RAT	O63358 rattus norv
425	5	16.7	962	1	PTRA_ECOLI	P05458 escherichia	498	5	16.7	2052	1	UBRB_SCHPO	O13731 schizosacch
426	5	16.7	965	1	PTO9_YEAST	P35252 saccharomyc	499	5	16.7	2114	1	MY9B_MOUSE	O9qy06 mus musculu
427	5	16.7	988	1	SY1_STNY3	P73505 synechocyst	500	5	16.7	2150	1	SDC3_CAEEL	P34706 caenorhabdi
428	5	16.7	1015	1	PPOL_BOVIN	P18493 bos taurus	501	5	16.7	2158	1	MY9B_HUMAN	Q13459 homo sapien
429	5	16.7	1036	1	YG35_YEAST	P35273 saccharomyc	502	5	16.7	2210	1	RRPL_EBOSM	P66802 ebola virus
430	5	16.7	1038	1	ITR4_HUMAN	P13612 homo sapien	503	5	16.7	2226	1	POLG_HPAV2	P26580 hepatitis a
431	5	16.7	1038	1	YKD3_YEAST	P36097 saccharomyc	504	5	16.7	2226	1	POLG_HPAV4	P26582 hepatitis a
432	5	16.7	1039	1	ITR4_MOUSE	O00651 mus musculu	505	5	16.7	2226	1	POLG_HPAV8	P26582 hepatitis a
433	5	16.7	1050	1	BU1B_HUMAN	O60566 homo sapien	506	5	16.7	2227	1	POLG_HPAVH	P08617 hepatitis a
434	5	16.7	1057	1	VP2_AHSV	O89508 african hor	507	5	16.7	2227	1	POLG_HPAVL	P06441 hepatitis a
435	5	16.7	1058	1	CARB_LACPL	P77886 lactobacilli	508	5	16.7	2227	1	POLG_HPAVM	P13901 hepatitis a
436	5	16.7	1070	1	DM1L_XENLA	P51892 xenopus lae	509	5	16.7	2230	1	POLG_HPAV2	P14553 simian hepa
437	5	16.7	1074	1	SM5A_HUMAN	O13591 homo sapien	510	5	16.7	2261	1	POLG_HPAV4	P26582 hepatitis a
438	5	16.7	1075	1	PST2_SCHPO	O13919 schizosacch	511	5	16.7	2298	1	CU05_HUMAN	O9y3r5 homo sapien
439	5	16.7	1077	1	SM5A_MOUSE	O62217 mus musculu	512	5	16.7	2326	1	PGG2_RAT	O00657 rattus norv
440	5	16.7	1078	1	S24A_HUMAN	O95486 homo sapien	513	5	16.7	2358	1	YEBJ_ECOLI	P76347 escherichia
441	5	16.7	1087	1	PGDS_XENLA	P26619 xenopus lae	514	5	16.7	2504	1	FAS_HUMAN	P49337 rattus norv
442	5	16.7	1088	1	RRPO_ROTBR	P17468 bovine rota	515	5	16.7	2505	1	PAS_RAT	P12785 rattus norv
443	5	16.7	1088	1	RRPO_ROTBR	P12615 bovine rota	516	5	16.7	2524	1	NOTC_XENLA	P21783 xenopus lae
444	5	16.7	1088	1	RRPO_ROTBR	P17699 porcine rot	517	5	16.7	2555	1	PPS3_BACSU	P39847 bacillus su
445	5	16.7	1088	1	RRPO_ROTBR	P22678 simian 11 r	518	5	16.7	2660	1	YEBJ_ECOS7	O8x8r7 escherichia
446	5	16.7	1116	1	VP2_RDVF	O98632 rice dwarf	519	5	16.7	2920	1	CLR2_MOUSE	O9r0m0 mus musculu
447	5	16.7	1116	1	YK54_AQUAE	O67838 aquilex aeo	520	5	16.7	2923	1	CLR2_HUMAN	O9hcu4 homo sapien
448	5	16.7	1131	1	AC15_MOUSE	P35601 mus musculu	521	5	16.7	3014	1	CLR1_HUMAN	O9hcu4 homo sapien
449	5	16.7	1133	1	VG1M_HANTB	P28728 hantaan vir	522	5	16.7	3034	1	CLR1_MOUSE	O35161 mus musculu
450	5	16.7	1133	1	VG1M_SEO8S	P13455 beoul vitus	523	5	16.7	3163	1	POLG_TUVVO	O20597 t genome po
451	5	16.7	1133	1	VG1M_SEO8S	P17880 beoul vitus	524	5	16.7	3164	1	POLG_TUVVU	P89509 t genome po
452	5	16.7	1134	1	VG1M_SEO8R	P28729 beoul vitus	525	5	16.7	3209	1	RELN_CHICK	O93574 gallus galli
453	5	16.7	1142	1	GLG1_CHICK	O02391 gallus galli	526	5	16.7	3317	1	CADN_RAT	P58363 rattus norv
454	5	16.7	1148	1	VP2_RDVO	O55519 rice dwarf	527	5	16.7	3321	1	PCN2_HUMAN	O95615 homo sapien
455	5	16.7	1160	1	GLG1_CRIGR	O9r1e9 cricetus	528	5	16.7	3354	1	CADN_HUMAN	O9hcu4 homo sapien
456	5	16.7	1161	1	POL_SFVL	P23074 simian foam	529	5	16.7	3354	1	CADN_MOUSE	O99p21 mus musculu
457	5	16.7	1169	1	EX5B_BORBU	O51578 borrelia bu	530	5	16.7	3358	1	PGCV_MOUSE	O62059 mus musculu
458	5	16.7	1171	1	GLG1_RAT	Q62638 ratulus norv	531	5	16.7	3579	1	STAN_DROME	O9v5t8 drosophila
459	5	16.7	1172	1	CU14_SCHPO	P41003 schizosacch	532	5	16.7	3587	1	SRF2_BACSU	O04747 bacillus su
460	5	16.7	1173	1	GIGA_ORYSA	O9xw17 oryza sativ	533	5	16.7	4344	1	DYHC_EMENT	P45444 emeritella
461	5	16.7	1174	1	DP3A_YERPE	O68770 yersinia pe	534	5	16.7	4349	1	FAT2_HUMAN	O9ny88 homo sapien
462	5	16.7	1174	1	KCRF_STRPU	P18294 strongyloce	535	5	16.7	4351	1	FAT2_RAT	O88277 rattus norv
463	5	16.7	1175	1	GLG1_MOUSE	Q61543 mus musculu	536	5	16.7	4351	1	TYCC_BACBR	O30409 b tyrocidin
464	5	16.7	1179	1	GLG1_HUMAN	Q9r896 homo sapien	537	5	16.7	4351	1	UC04_MAIZE	P80610 zea mays (m
465	5	16.7	1211	1	MP10L_HUMAN	O9bxc6 homo sapien	538	5	16.7	4351	1	ATPO_PIG	P80021 sus scrofa
466	5	16.7	1225	1	KP4A_CHICK	O90640 gallus galli	539	5	16.7	4351	1	CDK6_MOUSE	O64201 mus musculu
467	5	16.7	1230	1	SR20_CANAL	O9x212 candida alb	540	5	16.7	4351	1	NPF_MONEX	P41967 monelia ex
468	5	16.7	1268	1	S24B_HUMAN	O95487 homo sapien	541	5	16.7	4351	1	LHB3_RHOAC	P35093 rhodopsendo
469	5	16.7	1276	1	PWP6_CHLPP	O9r889 chlamydia p	542	5	16.7	4351	1	LHB5_RHOAC	P26790 rhodopsendo
470	5	16.7	1298	1	YTFN_HAEIN	O57523 haemophilus	543	5	16.7	4351	1	S108_BOVIN	P28782 bos taurus
471	5	16.7	1377	1	C1D_DROME	P19538 drosophila	544	5	16.7	4351	1	LCA_MACGI	P19122 macropus gi



545	4	13.3	46	1	OPT_BOVIN	P58874	bos	taurus	618	4	13.3	88	1	CX9A	CONTE	Q9gu58	canis	texti
546	4	13.3	47	1	YOFB_ECOLI	P76265	escherichia		619	4	13.3	88	1	VL1Q	HELIP	O25964	helicobacte	
547	4	13.3	50	1	COAB_BP22	P03618	bacterioph		620	4	13.3	88	1	VL1Q	VACCC	P21007	vaccinia vi	
548	4	13.3	52	1	ATP8_PENMO	Q9mgd7	penaeus	mon	621	4	13.3	88	1	VK03	VACCC	P20639	vaccinia vi	
549	4	13.3	52	1	LECA_VICSA	P16350	vicia	sativ	622	4	13.3	88	1	VK03	VACCV	P18378	vaccinia vi	
550	4	13.3	53	1	LECI_LATOC	P12306	lathyrus	oc	623	4	13.3	89	1	FL1Q	ECOLI	P33314	escherichia	
551	4	13.3	53	1	LEC2_LATOC	P12307	lathyrus	oc	624	4	13.3	89	1	FL1Q	ERWCA	P34201	erwinia car	
552	4	13.3	53	1	LECA_LATAP	P07441	lathyrus	ap	625	4	13.3	89	1	FL1Q	SALTU	P54701	salmonella	
553	4	13.3	53	1	LECA_VICCR	P07442	lathyrus	ar	626	4	13.3	89	1	GON3	PORNO	P51922	porcithyus	
554	4	13.3	53	1	LECA_VICCR	P02868	vicia	cracc	627	4	13.3	89	1	YAL1	BACLI	Q99142	bacillus li	
555	4	13.3	54	1	LECA_LATCI	P07440	lathyrus	ci	628	4	13.3	90	1	AF1Q	HUMAN	O13015	homo sapien	
556	4	13.3	54	1	LECA_LATCI	P07443	lathyrus	hi	629	4	13.3	90	1	AF1Q	MOUSE	P29783	mus musculu	
557	4	13.3	54	1	LECA_LATOD	P02869	lathyrus	od	630	4	13.3	90	1	GON3	DICLA	Q91a09	diceritarch	
558	4	13.3	55	1	LECA_LATCI	P07444	lathyrus	ci	631	4	13.3	90	1	GON3	PAGMA	P51921	pagrus majo	
559	4	13.3	55	1	LECA_LATSA	P12308	lathyrus	sa	632	4	13.3	90	1	GON3	SPAAU	P51923	sparus auro	
560	4	13.3	57	1	YCS7_HAEIN	P44143	haemophilus		633	4	13.3	90	1	IF1	MYCSP	P38037	mycoplasma	
561	4	13.3	58	1	YCS8_BPPT3	P03629	bacterioph		634	4	13.3	90	1	PAK2	SVNY3	P74554	mycobocyst	
562	4	13.3	59	1	RL30_STAMM	O06444	staphylococ		635	4	13.3	90	1	RPOL	SULAC	P46217	sulfolobus	
563	4	13.3	59	1	RL35_MYCCE	P47439	mycoplasma		636	4	13.3	90	1	RPOZ	STRCO	Q8Kx61	etereptomyce	
564	4	13.3	59	1	RL35_MYCEN	P75447	mycobacter		637	4	13.3	90	1	Z600	DROME	P22469	dirosophylla	
565	4	13.3	60	1	YRH1_AZOV1	Q44557	azotobacter		638	4	13.3	91	1	INAT	ECOLI	P19767	escherichia	
566	4	13.3	64	1	YCS3_ARCFU	O29015	archaeoglob		639	4	13.3	91	1	INSA	SHIFL	P19763	shigella fl	
567	4	13.3	66	1	ATNG_HUMAN	P54770	homo sapien		640	4	13.3	91	1	Y97	ADRO7	P05668	human adeno	
568	4	13.3	66	1	COPE_HBAPF	O32620	helicobacte		641	4	13.3	92	1	Y866	MERTH	O26972	methanobact	
569	4	13.3	66	1	KED2_ECOLI	P13967	escherichia		642	4	13.3	93	1	LSM5	YEAST	P40089	saccharomyc	
570	4	13.3	66	1	SECE_ARCFU	O29714	archaeoglob		643	4	13.3	93	1	Y435	METUA	P51019	homo sapien	
571	4	13.3	66	1	Y13G_BP74	P39500	bacterioph		644	4	13.3	93	1	Y435	METUA	O57877	methanococ	
572	4	13.3	67	1	RROZ_LISIN	O92a12	listeria in		645	4	13.3	93	1	YCBI	PSRDE	P29943	pseudomonas	
573	4	13.3	67	1	RROZ_LISMO	O89673	listeria mo		646	4	13.3	94	1	GATC	AQYAE	O67994	aquifex aeo	
574	4	13.3	67	1	YANI_YEAST	P39561	saccharomyc		647	4	13.3	94	1	RPCI	BP434	P61617	bacterioph	
575	4	13.3	70	1	CSPJ_SALTI	P58726	fremyella d		648	4	13.3	95	1	FXV6	HUMAN	Q9h0q3	homo sapien	
576	4	13.3	70	1	PYS2_FRBDI	P14878	fremyella d		649	4	13.3	95	1	NOLP	RHILP	P33717	rhizobium 1	
577	4	13.3	71	1	RROZ_CLOPE	O8Jk9	clostridium		650	4	13.3	95	1	RAG1	ORYSA	O01864	oryza sativ	
578	4	13.3	71	1	Y647_METJA	O58063	methanococ		651	4	13.3	95	1	RS14	CARRU	Q9a184	carsonella	
579	4	13.3	72	1	CHI_EBROB	P58845	procambarus		652	4	13.3	95	1	YBUC	ECOLI	P46119	escherichia	
580	4	13.3	72	1	HPTI_ECTVA	P38941	ectochlorho		653	4	13.3	96	1	CH10	AMOPS	P26005	amoeba prot	
581	4	13.3	72	1	RPOH_THEVO	Q97900	thermoplasm		654	4	13.3	96	1	CH10	CHRYI	P21205	chromatium	
582	4	13.3	72	1	VB11_VACCV	O01229	vaccinia vi		655	4	13.3	96	1	CH10	LEGEM	P26195	legionella	
583	4	13.3	73	1	COAB_BPFD	P03617	bacterioph		656	4	13.3	96	1	CH10	LEGEM	P26879	legionella	
584	4	13.3	73	1	SMS2_PLAAB	P21780	platicthyus		657	4	13.3	97	1	Y4D6	STRCO	P37977	streptomyce	
585	4	13.3	73	1	UH11_HSV7J	P53358	human herpes		658	4	13.3	97	1	YCIS	HAEIN	P44129	haemophilus	
586	4	13.3	73	1	Y43_BPRI8	P18243	bacterioph		659	4	13.3	98	1	ELIA	PHYCP	P15571	phytophthor	
587	4	13.3	74	1	ATPL_RICCN	Q92Jp1	rickettsia		660	4	13.3	98	1	ELIA	PHYME	P35658	phytothor	
588	4	13.3	74	1	MIH_PROBO	Q10987	procambarus		661	4	13.3	98	1	NULM	SIGHI	O15657	sigmodon hi	
589	4	13.3	74	1	SMS2_MYOSC	P09876	tyrocephala		662	4	13.3	99	1	BXAL	SAMCY	P33718	samia cyth	
590	4	13.3	74	1	Y777_TREPA	O83756	treponema p		663	4	13.3	99	1	YG42	CAMTE	O9pm74	samia cyth	
591	4	13.3	75	1	GLUC_AMICA	P33528	ania calva		664	4	13.3	99	1	YO10	BEHPI	P51712	bacterioph	
592	4	13.3	75	1	MCMB_ECOLI	Q9rm53	escherichia		665	4	13.3	100	1	BXAL	SAMCY	P33719	samia cyth	
593	4	13.3	75	1	WC12_SPVKA	P32221	swinepox vi		666	4	13.3	100	1	CHLB	PLRSC	P37854	pleurozium	
594	4	13.3	75	1	YODD_ECOLI	P76328	escherichia		667	4	13.3	100	1	CHLB	POLCU	P37852	polycrithum	
595	4	13.3	75	1	YRKI_BACSU	P54436	bacillus su		668	4	13.3	100	1	ES6Y	MYCLE	Q49945	mycobacteri	
596	4	13.3	77	1	VG43_BPMU5	O05255	mycobacteri		669	4	13.3	100	1	SECG	RICPR	O92668	rickettsia	
597	4	13.3	77	1	YDHI_PLAAS	P14587	plasmodium		670	4	13.3	100	1	VE7	HPV22	P50760	methanococ	
598	4	13.3	79	1	SPDI_ARABI	P46802	araneus bic		671	4	13.3	100	1	Y953	METUA	P50760	methanococ	
599	4	13.3	80	1	IF1_DEIRA	Q9rsk1	deinococcus		672	4	13.3	100	1	YIS1	SHISO	P16939	shigella so	
600	4	13.3	80	1	PSAC_SKECO	O96804	skeletoenema		673	4	13.3	101	1	DMT1	ALIMI	O9pu60	alligator m	
601	4	13.3	81	1	PSAC_ODOSI	P49477	odontella s		674	4	13.3	101	1	NIFU	ANAVA	O53349	anabena va	
602	4	13.3	82	1	CZ_OXYNO	P05526	oxytricha n		675	4	13.3	101	1	SMO3	YEAST	P43331	saccharomyc	
603	4	13.3	83	1	NI9M_BOVIN	Q02371	bos caurus		676	4	13.3	101	1	VG7	BPB03	Q37867	bacterioph	
604	4	13.3	83	1	PGCA_PIG	Q29011	bos scrofa		677	4	13.3	101	1	YH65	YEAST	P37881	saccharomyc	
605	4	13.3	83	1	PSAK_SYNN	P20453	synechococ		678	4	13.3	101	1	YH65	YEAST	Q40481	saccharomyc	
606	4	13.3	83	1	PROK_SULAC	P30463	sulfolobus		679	4	13.3	102	1	ANTR	RANCA	P00111	rana catesb	
607	4	13.3	84	1	CPPA_NEIGO	P07048	neisseria g		680	4	13.3	102	1	GP4D	CHIMU	Q64440	chlamydia m	
608	4	13.3	85	1	CYC6_PLEBO	P00117	plectonema		681	4	13.3	102	1	GP4D	CHIPS	Q46262	chlamydia p	
609	4	13.3	85	1	PSAK_SYNTU	P23318	synechococ		682	4	13.3	102	1	GP4D	CHIRP	P08785	chlamydia t	
610	4	13.3	85	1	YPBS_BACSU	P54160	bacillus su		683	4	13.3	102	1	RP5W	PSEPU	P15552	pseudomonas	
611	4	13.3	86	1	OXIR_PIG	O97661	bos scrofa		684	4	13.3	102	1	VE7	HPV20	P05078	human papil	
612	4	13.3	86	1	PSAK_PORPU	P51370	porphyra pu		685	4	13.3	103	1	ACPH	CLOAB	Q97h14	clostridium	
613	4	13.3	87	1	FXV4_RAT	O63113	rattus norv		686	4	13.3	103	1	CHLB	ARABE	P37843	araucaria h	
614	4	13.3	87	1	PSAK_GUTH	O74444	gulliardia		687	4	13.3	103	1	CHLB	BATTR	P37844	bazzania tr	
615	4	13.3	87	1	VP08_BPAPS	Q991u0	bacterioph		688	4	13.3	103	1	CHLB	BOUAR	P37845	equisetum a	
616	4	13.3	87	1	Y04A_BP74	P07083	bacterioph		689	4	13.3	103	1	CHLB	BOUSC	Q32214	equisetum s	
617	4	13.3	87	1	YBBD_ECOLI	P30977	escherichia		690	4	13.3	103	1	CHLB	ISOEC	Q32450	isoetes ech	

691	4	13.3	103	1	CHLB_LYCANA	P37847 lycopodium	764	4	13.3	115	1	HYPA_AQUAE	O67133 aquilex aeo
692	4	13.3	103	1	CHLB_LYCOCO	P37848 lycopodium	765	4	13.3	115	1	RK22_ASTRLO	P34770 aetasia lon
693	4	13.3	103	1	CHLB_METECO	P37849 metasequoia	766	4	13.3	115	1	RL24_SYNY3	P73304 synecocyst
694	4	13.3	103	1	CHLB_NEPX	P37850 nephrrolepis	767	4	13.3	115	1	SMS1_PROAN	O9w760 protopneus
695	4	13.3	103	1	CHLB_OSMCL	P37851 osmundia cia	768	4	13.3	115	1	YBR2_YEAST	P38229 saccharomyc
696	4	13.3	103	1	CHLB_SELMO	P37856 selaginella	769	4	13.3	115	1	YD30_HELPY	O25888 helicobacte
697	4	13.3	103	1	CHLB_ZAMFI	P37857 zamia flesch	770	4	13.3	115	1	YL22_CABEL	P24412 caenorhabdi
698	4	13.3	103	1	HE2_HUMAN	O08648 homo sapien	771	4	13.3	116	1	RPBY_ARATH	O38859 arabidopsis
699	4	13.3	103	1	PRO2_STRPN	O97pa4 streptococc	772	4	13.3	116	1	RT13_DAVCA	P23209 daucus caro
700	4	13.3	103	1	S112_HUMAN	O96fc6 homo sapien	773	4	13.3	116	1	RT13_TOBAC	P05488 nicotiana t
701	4	13.3	103	1	YPA1_LEGPN	P26880 legionella	774	4	13.3	116	1	RT13_WHEAT	P072924 triticum ae
702	4	13.3	104	1	YPC6_CLAPU	P22368 claviaceps p	775	4	13.3	116	1	Y828_PYRHO	O58558 pyrococcus
703	4	13.3	104	1	NEUY_LAMFL	P48097 lampetra fl	776	4	13.3	116	1	Y157_AQUAE	O67709 aquilex aeo
704	4	13.3	104	1	RL30_KLULA	P38664 kluyveromyc	777	4	13.3	117	1	RBS1_HYDMR	O59459 hydrogenvol
705	4	13.3	104	1	PRO2_STRPY	P82577 streptococc	778	4	13.3	117	1	RPOZ_IACIA	O59459 lactococcus
706	4	13.3	104	1	THIO_BOVIN	O97680 bos taurus	779	4	13.3	118	1	ELI1_PHYCR	P44824 phytophthor
707	4	13.3	104	1	THIO_CALTA	O9bdj3 callithrix	780	4	13.3	118	1	KDGL_HA1RN	P44824 haemophilus
708	4	13.3	104	1	THIO_CHICK	O08629 gallus gall	781	4	13.3	118	1	PAND_CAUOR	O9a602 caulobacter
709	4	13.3	104	1	THIO_HORSE	O97508 equus cabal	782	4	13.3	118	1	VNS2_IATRS	O82729 influenza a
710	4	13.3	104	1	THIO_HUMAN	P10599 homo sapien	783	4	13.3	118	1	YD30_HELPY	O9zj94 helicobacte
711	4	13.3	104	1	THIO_MACMU	P29451 macaca mula	784	4	13.3	118	1	YUSI_BACSU	O32175 bacillus su
712	4	13.3	104	1	THIO_MOUSE	P10639 mus musculu	785	4	13.3	119	1	AMCY_MEREX	P04172 methyllobact
713	4	13.3	104	1	THIO_OPFHA	O98tx1 ophiophagus	786	4	13.3	119	1	RL19_MYCPN	P75133 mycoplasma
714	4	13.3	104	1	THIO_PIG	P82460 sus scrofa	787	4	13.3	119	1	VATF_HUMAN	O16864 homo sapien
715	4	13.3	104	1	THIO_PABIT	P08628 oryctolagus	788	4	13.3	119	1	VATF_RAT	P50408 rattus norv
716	4	13.3	104	1	THIO_SHEEP	P11232 rattus norv	789	4	13.3	119	1	VP26_HYSEB	P28974 equine herp
717	4	13.3	104	1	VER7_HPV26	P50413 ovis aries	790	4	13.3	120	1	PAND_RALSO	O8xv66 ralestonia s
718	4	13.3	104	1	VER7_HPV26	P36824 human papil	791	4	13.3	120	1	RL7A_HALNI	O9h088 halobacteri
719	4	13.3	104	1	YSH7_DICDI	P13327 diacyosteli	792	4	13.3	120	1	YACL_ECOLI	P45557 escherichia
720	4	13.3	104	1	Y907_HA1RN	P44072 haemophilus	793	4	13.3	120	1	YJY8_YEAST	P47091 saccharomyc
721	4	13.3	105	1	N1FW_ANASP	O44149 anabaena sp	794	4	13.3	121	1	FLIO_ECOLI	P22566 escherichia
722	4	13.3	105	1	SMS2_ICTPU	P01172 ictalurus p	795	4	13.3	121	1	LCA_MACRG	P07458 macrocypus ru
723	4	13.3	106	1	PUTX_PSEBU	P00259 pseudomonas	796	4	13.3	121	1	PFDB_MERTH	O26774 methanobact
724	4	13.3	106	1	SRCE_HA1RN	P43805 haemophilus	797	4	13.3	121	1	RBS_ALVHS	P24682 alvinococh
725	4	13.3	106	1	Y1P1_VIBCH	O9Kk29 vibrio chol	798	4	13.3	121	1	TTTH_CABEL	O21882 caenorhabdi
726	4	13.3	106	1	Y368_METUA	O57814 methanococc	799	4	13.3	122	1	HEX3_ADETI	P04376 lupata aden
727	4	13.3	106	1	Y488_ARCFU	O29762 archaeoglob	800	4	13.3	122	1	S109_BOVIN	P28783 bos taurus
728	4	13.3	107	1	IFLC_MAIZE	P46616 zea mays (m	801	4	13.3	123	1	CRCB_PYRAB	O5v0x2 pyrococcus
729	4	13.3	107	1	IFLC_ORYSA	P13135 oryza sativ	802	4	13.3	123	1	CRCB_PYRAB	O59171 pyrococcus
730	4	13.3	107	1	PPDC_ECOLI	P08372 escherichia	803	4	13.3	123	1	KADI_RAT	P39069 rattus norv
731	4	13.3	107	1	VNS1_PVM	P28888 pneumonia v	804	4	13.3	123	1	LCA_PAREY	P12065 papio cynoc
732	4	13.3	107	1	YFHP_HA1RN	P46672 haemophilus	805	4	13.3	123	1	LSM6_YEAST	O06406 saccharomyc
733	4	13.3	107	1	YNIU_AZOVI	O44540 azotobacter	806	4	13.3	123	1	NB4M_NEUCR	P42114 neurospora
734	4	13.3	108	1	GLRX_YEAST	P17695 saccharomyc	807	4	13.3	123	1	RL7A_MERYA	O8t0v3 methanopyru
735	4	13.3	108	1	NIGM_BOVIN	O02374 bos taurus	808	4	13.3	123	1	YQ30_MYCLE	O06091 mycobacteri
736	4	13.3	108	1	PT91_YEAST	O02772 saccharomyc	809	4	13.3	124	1	AT14_YEAST	O12349 saccharomyc
737	4	13.3	108	1	VAPI_BACNO	O45560 bacteroides	810	4	13.3	124	1	YABI_BACSU	B37552 bacillus su
738	4	13.3	109	1	C550_NITWI	P00085 nitrobacter	811	4	13.3	124	1	YMO2_CABEL	P34493 caenorhabdi
739	4	13.3	109	1	FEI2_FELCA	P30440 felis silve	812	4	13.3	125	1	CUI4_MANSE	P13229 manduca sex
740	4	13.3	109	1	PRVA_HUMAN	P20472 homo sapien	813	4	13.3	125	1	FABL_HALBI	P81653 halielurus
741	4	13.3	109	1	VMTM_LAMBD	P03737 bacteriopho	814	4	13.3	125	1	FABL_RHSA	P80856 rhandia sap
742	4	13.3	109	1	YR1F_ECOLI	O46953 escherichia	815	4	13.3	125	1	MINK_CAYPO	O60409 cavia porce
743	4	13.3	110	1	FER_ECOLI	P25528 escherichia	816	4	13.3	125	1	REC9_SCHPO	O09177 schizosacch
744	4	13.3	110	1	GBG_YEAST	P18852 saccharomyc	817	4	13.3	125	1	SECG_BORBU	O51063 borrelia bu
745	4	13.3	110	1	H1S2_CLOAB	O97Kh6 clostridium	818	4	13.3	125	1	SM52_LOPAM	O01170 lophus ame
746	4	13.3	110	1	NJOK_RICCN	O92G98 rickettsia	819	4	13.3	125	1	VG61_BPM15	O05274 mycobacteri
747	4	13.3	110	1	NJOK_RICPR	O92G98 rickettsia	820	4	13.3	125	1	Y568_MERYA	O57968 methanococ
748	4	13.3	110	1	VATF_BOVIN	O28029 bos taurus	821	4	13.3	125	1	Y811_SUITO	O97376 sulfolobus
749	4	13.3	110	1	YBSA_MERYA	P81316 methanococc	822	4	13.3	125	1	Y944_HELPY	O25598 helicobacte
750	4	13.3	111	1	IMWI_ECOLI	P08701 escherichia	823	4	13.3	126	1	ALDR_LACLA	O34143 lactococcus
751	4	13.3	112	1	NBR2_HUMAN	O15453 homo sapien	824	4	13.3	126	1	PAND_PSEAE	O94v88 pseudomonas
752	4	13.3	112	1	RBAFA_MYCGE	P47389 mycoplasma	825	4	13.3	126	1	Y681_MYCPU	O50355 mycoplasma
753	4	13.3	112	1	RS6_CHLPN	O96v65 chlamydia p	826	4	13.3	126	1	YA42_SCHPO	O09723 schizosacch
754	4	13.3	112	1	Y112_ADBO2	P03289 human adeno	827	4	13.3	127	1	YMO6_SUITO	O97u19 sulfolobus
755	4	13.3	112	1	YCO1_AQUAE	O61257 aquilex aeo	828	4	13.3	127	1	CRCB_THEMA	O9xv88 thermotoga
756	4	13.3	113	1	IF1C_WHEAT	P58272 triticum ae	829	4	13.3	127	1	RT12_CHOCR	P48858 chondrus cr
757	4	13.3	113	1	RL24_SYNP6	O24700 synecococc	830	4	13.3	127	1	TRM1_ECOLI	P10026 escherichia
758	4	13.3	114	1	D2_ONCVO	P54187 onchocerca	831	4	13.3	127	1	TRM6_ECOLI	P18808 escherichia
759	4	13.3	114	1	DCHS_LACHU	P04193 lactobacilli	832	4	13.3	127	1	TRM7_ECOLI	P33787 escherichia
760	4	13.3	114	1	GVJ2_STRCO	O97j58 streptomyce	833	4	13.3	127	1	YUGF_ECOLI	P39330 escherichia
761	4	13.3	114	1	RT13_OENBE	P15758 oenothera b	834	4	13.3	127	1	YL35_ARCFU	O2815 archaeoglob
762	4	13.3	114	1	Y941_ARCFU	O23221 archaeoglob	835	4	13.3	127	1	YVNI_AZOVI	P40431 azotobacter
763	4	13.3	115	1	ANFC_TRISC	P55208 tritakis scy	836	4	13.3	128	1	RS9_THETH	P80374 thermus the

837	4	13.3	128	1	VATF_ARATH	Q9ZQ4 arabidopeis	910	4	13.3	137	1	ATPE_MEDSA	Q9TK6 medicago sa
838	4	13.3	128	1	V371_BUCAI	P57452 buchiera ap	911	4	13.3	137	1	ATPE_ORYSA	P12066 oryza sativ
839	4	13.3	128	1	YCDK_ECOLI	P5896 escherichia	912	4	13.3	137	1	ATPE_PEA	P05039 pisum sativ
840	4	13.3	129	1	COXE_YEAST	P32759 saccharomyc	913	4	13.3	137	1	ATPE_PICAB	Q47036 picea abies
841	4	13.3	129	1	HMFI_YEAST	P40073 saccharomyc	914	4	13.3	137	1	ATPE_PINTH	P41633 pinus thunb
842	4	13.3	129	1	MINK_MOUSE	P23299 mus musculu	915	4	13.3	137	1	ATPE_MHEAT	P20859 triticum ae
843	4	13.3	129	1	RHCB_AGRKH	P81398 agkistrodon	916	4	13.3	137	1	CHH_PROCL	Q25663 procambiar
844	4	13.3	129	1	RS9_ECOLI	P02363 escherichia	917	4	13.3	137	1	FLGB_ECOLI	P75934 escherichia
845	4	13.3	129	1	RS9_HAEIN	P44388 haemophilus	918	4	13.3	137	1	IGJ_HUMAN	P01591 homo sapien
846	4	13.3	129	1	RS9_HAESO	P31782 haemophilus	919	4	13.3	137	1	URE2_STAXY	P42874 staphylococ
847	4	13.3	129	1	RS9_SALTY	ORFX5 salmonella	920	4	13.3	137	1	Y890_MENJA	Q58308 methanococ
848	4	13.3	129	1	RT13_MAIZE	P08977 zea mays (m	921	4	13.3	137	1	YAB3_PSEPU	P45388 pseudomonas
849	4	13.3	129	1	TDCF_ECOLI	P42631 escherichia	922	4	13.3	138	1	ATPE_BUCAI	P57153 buchiera ap
850	4	13.3	129	1	YBAZ_ECOLI	P75707 escherichia	923	4	13.3	138	1	ATPE_BUCAP	P51871 buchiera ap
851	4	13.3	129	1	YD39_HASIN	P71379 haemophilus	924	4	13.3	138	1	FLGN_ECOLI	P43533 escherichia
852	4	13.3	129	1	YEB6_PASMU	Q91656 pasteurella	925	4	13.3	138	1	HUPG_RHOCA	Q03006 rhodobacter
853	4	13.3	130	1	DHSD_HALNI	Q9h652 halobacteri	926	4	13.3	138	1	NUSB_HELPU	Q9ZNS7 helicobacte
854	4	13.3	130	1	DHSD_NATPH	P72109 natriomonas	927	4	13.3	138	1	NUSB_HELPU	Q24853 helicobacte
855	4	13.3	130	1	DSRE_CHRVT	Q87896 chromatium	928	4	13.3	138	1	SUFE_ECOLI	P76194 salmonella
856	4	13.3	130	1	MINK_PIG	Q9TUN9 sus scrofa	929	4	13.3	139	1	YTL1_SALTY	Q9f752 petunia hyb
857	4	13.3	130	1	MINK_RABIT	Q28705 oryctolagus	930	4	13.3	139	1	ADP1_PETRY	Q41764 zea mays (m
858	4	13.3	130	1	MINK_RAT	Q97708 methanococ	931	4	13.3	139	1	ADP3_MAIZE	Q8X158 clostridium
859	4	13.3	130	1	RS8_METTL	P14038 methanococ	932	4	13.3	139	1	PABZ_CLOPE	P09638 alcaligenes
860	4	13.3	130	1	RS8_METVA	Q97703 methanococ	933	4	13.3	139	1	RBS_CALCEU	P09638 alcaligenes
861	4	13.3	130	1	RS8_METVO	P57470 buchiera ap	934	4	13.3	139	1	RBS_THAMO	Q98948 thalassiosi
862	4	13.3	130	1	RS9_BUCAI	Q9ENB1 pasteurella	935	4	13.3	139	1	RK16_PORPU	P51307 porphyra pu
863	4	13.3	130	1	RS9_PASMU	Q9KUF2 vibrio chol	936	4	13.3	139	1	RL16_SVNY3	P73133 synechocyst
864	4	13.3	130	1	RS9_VIBCH	Q8Z62 yersinia pe	937	4	13.3	139	1	Y252_AOUAE	Q666132 aquifex aeo
865	4	13.3	130	1	RS9_YERBE	Q10328 oryglia pseu	938	4	13.3	140	1	EXB1_XANCP	Q34259 xanthomonas
866	4	13.3	130	1	Y075_NPPOP	Q23536 caenorhabdi	939	4	13.3	140	1	FLGN_SALTY	P37406 salmonella
867	4	13.3	130	1	YHM6_CABEL	P44839 haemophilus	940	4	13.3	140	1	LCA_MACEU	Q06653 macropus eu
868	4	13.3	130	1	YUGF_HASIN	Q52178 myxococcus	941	4	13.3	140	1	RS19_SUTSO	Q06653 macropus eu
869	4	13.3	131	1	DFRA_MYXXA	Q8Z329 pyrobaculum	942	4	13.3	140	1	Y128_LISMO	Q9UX33 sulfobius
870	4	13.3	131	1	HIS3_PYRAE	Q8Z329 pyrobaculum	943	4	13.3	140	1	Y175_LISIN	P58702 listeria mo
871	4	13.3	131	1	Y624_MYCTU	P96914 mycobacteri	944	4	13.3	140	1	YAB3_SCHPO	Q92F01 listeria in
872	4	13.3	132	1	AGSR_HUMAN	C00253 homo sapien	945	4	13.3	140	1	YAB3_ECOLI	P009802 escherichia
873	4	13.3	132	1	ATPE_ARATH	P09468 arabidopsis	946	4	13.3	140	1	YEB6_MYCPN	P40771 escherichia
874	4	13.3	132	1	CYCP_PARSP	P00143 paracoccus	947	4	13.3	141	1	ATPE_PSEAE	P75317 mycoplasma
875	4	13.3	133	1	ACPS_CLOPE	Q8XNP1 clostridium	948	4	13.3	141	1	HBA_TRLIN	Q9HLC1 pseudomonas
876	4	13.3	133	1	AMPE_RAT	P50123 rattus norv	949	4	13.3	141	1	HBFI_URECA	P06148 urechis cau
877	4	13.3	133	1	ATPE_TROBA	P07138 ipomea bat	950	4	13.3	141	1	NIFD_RHLIT	P06445 urechis cau
878	4	13.3	133	1	ATPE_OENHO	Q9MCP9 oenothera h	951	4	13.3	141	1	VE6_HPV17	P36805 human papil
879	4	13.3	133	1	ATPE_TOSAC	P00834 nicotiana t	952	4	13.3	141	1	YJEB_ECOLI	P21498 escherichte
880	4	13.3	133	1	YCGK_ECOLI	P76002 escherichia	953	4	13.3	141	1	YORE_TTV1	P19289 thermoprote
881	4	13.3	133	1	YDUF_SCHPO	P87054 schizosacch	954	4	13.3	141	1	YROP_BACSU	P54192 bacillus su
882	4	13.3	133	1	YUFI_PLRBO	Q00247 pleconema	955	4	13.3	142	1	AGNX_APIPS	P12284 aplysia sp.
883	4	13.3	133	1	YZ01_AOUAE	Q66396 aquifex aeo	956	4	13.3	142	1	ASPI_STRTR	P80445 streptococ
884	4	13.3	134	1	ATPE_SPTOL	P00883 epinacia ol	957	4	13.3	142	1	ASPD_STRTR	Q30851 streptococ
885	4	13.3	134	1	D3_ONCVO	P54188 onchocerca	958	4	13.3	142	1	CYTD_HUMAN	P28325 homo sapien
886	4	13.3	134	1	PS11_SVNY3	P74367 pinus thunb	959	4	13.3	142	1	LCA_BOVIN	P00711 bos taurus
887	4	13.3	134	1	RK16_PINTH	P52767 pinus thunb	960	4	13.3	142	1	LCA_CAPRI	P00712 capra hircu
888	4	13.3	134	1	YD89_MENJA	Q58784 methanococ	961	4	13.3	142	1	LCA_HUMAN	P00709 homo sapien
889	4	13.3	134	1	YHCB_ECOLI	P39436 escherichia	962	4	13.3	142	1	LCA_SHEEP	P09462 ovis aries
890	4	13.3	135	1	ATPE_WARPO	P06285 marcanthia	963	4	13.3	142	1	PFDA_METAC	Q8C116 methanosarc
891	4	13.3	135	1	CHH1_ORCLI	Q25588 orconectes	964	4	13.3	142	1	RECK_PSEPU	P37862 pseudomonas
892	4	13.3	135	1	CHH2_ORCLI	P47929 homo sapien	965	4	13.3	142	1	VE6_HPV60	Q80941 human papil
893	4	13.3	135	1	LEG7_HUMAN	Q59384 escherichia	966	4	13.3	142	1	Y742_CHLPP	Q92733 chlamydia p
894	4	13.3	135	1	IGUL_ECOLI	Q60003 salmonella	967	4	13.3	142	1	YAL6_HAEIN	P44095 haemophilus
895	4	13.3	135	1	IGUL_SALTY	Q60003 salmonella	968	4	13.3	142	1	YE72_SCHPO	O14155 schistosach
896	4	13.3	135	1	REV_EIAV9	P11305 equine infe	969	4	13.3	142	1	YKX5_YEAST	P50943 saccharomyc
897	4	13.3	135	1	REV_EIAVC	P32543 equine infe	970	4	13.3	143	1	SM16_SCHMA	Q00716 schistosoma
898	4	13.3	135	1	RK16_EUGGR	P21512 euglena gra	971	4	13.3	144	1	YA72_AOUAE	Q67167 aquifex aeo
899	4	13.3	135	1	YOHV_BACSU	P54519 bacillus su	972	4	13.3	145	1	COPY_ENTHR	Q47839 enterococu
900	4	13.3	136	1	ATPE_CUSRE	P30400 cuscutea ref	973	4	13.3	145	1	DUT_CHIMU	O9PK82 chlamydia m
901	4	13.3	136	1	CALI_ARATH	P25854 arabidopsis	974	4	13.3	145	1	DUT_CHLTR	O84294 chlamydia t
902	4	13.3	136	1	CALI_ONCKE	P01263 oncorhynch	975	4	13.3	145	1	MMF1_YEAST	P40185 saccharomyc
903	4	13.3	136	1	NNDI_ECOLI	Q47415 escherichia	976	4	13.3	145	1	VP3_BPCHP	P19194 bacteriophage
904	4	13.3	136	1	NNDI_SALTI	Q84266 salmonella	977	4	13.3	146	1	AROF_STRLI	P55911 streptomyce
905	4	13.3	136	1	NNDI_SALTY	O56109 salmonella	978	4	13.3	146	1	PABZ_CAMJE	O9P162 campylobact
906	4	13.3	136	1	RS9_BORBU	O51313 borrelia bu	979	4	13.3	146	1	PABZ_PSEAE	O9HX72 pseudomonas
907	4	13.3	136	1	UK14_RAT	P52759 rattus norv	980	4	13.3	146	1	HBBI_ANAMT	P83272 anathichas
908	4	13.3	137	1	ACTP_ACACA	P37167 acanthamoeb	981	4	13.3	146	1	HBBI_PAGBO	O93348 pagothenia
909	4	13.3	137	1	ATPE_MAIZE	P00835 zea mays (m	982	4	13.3	146	1	HBBI_GADWO	O13077 gadus morhu

```

983 4 13.3 146 1 HBB_MELAE 009232 melangramm
984 4 13.3 146 1 HBB_MERMR 013078 melangius
985 4 13.3 146 1 MSNB_CLOAB 097100 clostridium
986 4 13.3 146 1 RL15_METTU 026133 methanobact
987 4 13.3 146 1 RL15_METTU P95071 mycobacteri
988 4 13.3 146 1 SMD1_YEAST Q02260 saccharomyc
989 4 13.3 146 1 SP22_BACSV P10728 bacillus su
990 4 13.3 146 1 YD34_MYCTU Q10645 mycobacteri
991 4 13.3 147 1 ARGGR_CHLPPN Q94821 chlamydia p
992 4 13.3 147 1 ARQO_PSEAE Q10557 pseudomonas
993 4 13.3 147 1 CALM_PASGY Q39752 fagus sylv
994 4 13.3 147 1 ENPP_BPT3 P10305 bacterioph
995 4 13.3 148 1 CAL1_PETRY P27162 petunia hyb
996 4 13.3 148 1 CAL2_ARATH P25069 arabidopsis
997 4 13.3 148 1 CAL4_ARATH Q03510 arabidopsis
998 4 13.3 148 1 CAL6_ARATH Q03509 arabidopsis
999 4 13.3 148 1 CALM_ACHTL P15094 achlya kleb
1000 4 13.3 148 1 CALM_BLAEM Q9hfy6 blastoclad

```

## ALIGNMENTS

RESULT 1  
SODM\_CANAL STANDARD; PRT; 234 AA.

```

AC 013401;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Superoxide dismutase [Mn], mitochondrial precursor (EC 1.15.1.1).
GN SOD2.
OS Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mtosporic Saccharomycetales; Candida.
ON NCBI_Taxid=5476;
RX MEDLINE=99177423; PubMed=10076057;
RA Rhee G.E., Huang C.S., Brady M.J., Kim S.T., Kim Y.R., Huh W.K.,
RA Baek Y.U., Lee B.H., Lee J.S., Kang S.O.;
RA "Manganese-containing superoxide dismutase and its gene from Candida
RT albicans".
RL Biochim. Biophys. Acta 1426:409-419(1999).
CC -1- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems.
CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -1- COFACTOR: Manganese.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; AF031478; AAB86583.1; -
DR HSSP; P04179; IABM.
DR InterPro; IPR001169; SODismutase.
DR Pfam; PF00081; sode; 1.
DR PROSITE; PS000475; SODismutase; 1.
DR PROSITE; PS00089; SOD NM; 1.
KW Oxidoreductase; Manganese; Mitochondrion; Transit peptide.
FT TRANSIT 1 34 MITOCHONDRION.
FT CHAIN 35 234 SUPEROXIDE DISMUTASE [MN].
FT METAL 60 60 MANGANESE (BY SIMILARITY).
FT METAL 108 108 MANGANESE (BY SIMILARITY).

```

```

FT METAL 198 198 MANGANESE (BY SIMILARITY).
FT METAL 202 202 MANGANESE (BY SIMILARITY).
SQ SEQUENCE 234 AA; 26173 MW; EBFPC2D769C1D9C1 CRC64;
Query Match 83.3%; Score 25; DB 1; Length 234;
Best Local Similarity 100.0%; Pred. No. 1.9e-19;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 KYSLPELDYFSATPEYISGOINEL 25
DB 35 KYSLPELDYFSATPEYISGOINEL 59

```

## RESULT 2

SODM\_YEAST STANDARD; PRT; 233 AA.

```

AC P00447;
DT 21-JUL-1986 (Rel. 01, Created)
DT 23-OCT-1986 (Rel. 02, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Superoxide dismutase [Mn], mitochondrial precursor (EC 1.15.1.1).
GN SOD2 OR YHR008C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycos.
ON NCBI_Taxid=4932;
RX MEDLINE=65127011; PubMed=3882422;
RA Mares C.A.M., van Loon A.P.G.M., Oudshoorn P., van Steeg H.,
RA Grivell L.A., Slater E.C.;
RA "Nucleotide sequence analysis of the nuclear gene coding for
RT manganese superoxide dismutase of yeast mitochondria, a gene
RT previously assumed to code for the Rieske iron-sulphur protein."
RL Eur. J. Biochem. 147:153-161(1985).

```

## SEQUENCE FROM N.A.

```

RX MEDLINE=94378003; PubMed=8091229;
RA Johnson M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
RA Du Z., Faveiro A., Fulton L., Gattung S., Gelsel C., Kirsten J.,
RA Knudsen T., Hillier L., Jier M., Johnston L., Langston Y.,
RA Latreille P., Louis E.J., Macri C., Marais E., Meneses S., Mouser L.,
RA Nham M., Rifkin L., Riles L., St Peter H., Trevaaskis E., Vaughan K.,
RA Vignati D., Wilcox L., Wohlman P., Waterson R., Wilson R.,
RA Vandin M.;
RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
RT VII."
RT Science 265:2077-2082(1994).
RN [3]
RP SEQUENCE OF 1-39 FROM N.A.
RX MEDLINE=89211942; PubMed=3072251;
RA Schranx I.S., Sims P.F., Oliver S.G.;
RT "Functional expression of the yeast Mn-superoxide dismutase gene in
RT Escherichia coli requires deletion of the signal peptide sequence."
RL Gene 73:121-130(1988).
RN [4]
RP SEQUENCE OF 27-233.
RA Ditlev C., Johansen J.T., Martin B.M., Svendsen I.;
RT "The complete amino acid sequence of manganese-superoxide dismutase
RT from Saccharomyces cerevisiae."
RL Carlsberg Res. Commun. 47:81-91(1982).
CC -1- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems.
CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -1- COFACTOR: Manganese.
CC -1- SUBUNIT: Homotetramer.
CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL; X02156; CAA26092.1; -;  
DR EMBL; U10400; AAB68939.1; -;  
DR EMBL; M24079; AAA35065.1; -;  
DR PIR; A00521; DSBYN.  
DR PIR; S46785; S46785.  
DR HSP; P04179; IABM.  
DR SGD; S0001050; SOD2.  
DR InterPro; IPR01189; SODismutase.  
DR Pfam; PF00081; sode; 1.  
DR Pfam; PF02777; sode C; 1.  
DR ProDom; PD000475; SODismutase; 1.  
DR PROSITE; PS00088; SOD\_MN; 1.  
DR Oxidoreductase; Manganese; Mitochondrion; Transist peptide.  
TRANSIT 1 26  
CHAIN 27 233 MITOCHONDRION.  
FT METAL 52 52 SUPEROXIDE DISMUTASE [MN].  
FT METAL 107 107 MANGANESE (BY SIMILARITY).  
FT METAL 194 194 MANGANESE (BY SIMILARITY).  
FT METAL 198 198 MANGANESE (BY SIMILARITY).  
SQ SEQUENCE 233 AA; 25774 MW; 88A9391FBB31D06B CRC64;

Query Match 33.3%; Score 10; DB 1; Length 233;  
Best Local Similarity 100.0%; Pred. No. 0.0012;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 EPIISQINE 24  
Db 41 EPIISQINE 50

RESULT 3  
SODM\_MYCAV STANDARD; PRT; 206 AA.  
ID SODM\_MYCAV  
AC P47201;

DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Superoxide dismutase [Mn] (EC 1.15.1.1).  
GN SODA OR SOD.

OS Mycobacterium avium.  
CC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;  
CC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
NCBI\_TaxID=1764;

-RN  
X1 SEQUENCE FROM N.A.  
RC STRAIN=TMC 724;  
RX MEDLINE=96276149; PubMed=8692009;  
RA Becuyer V.E., Haddad N., Frehel C., Berche P.;  
RT "Molecular characterization of a surface-exposed superoxide dismutase  
of Mycobacterium avium."  
RL Microb. Pathog. 20:41-55(1996).

-1- FUNCTION: Destroys radicals which are normally produced within the  
CC cells and which are toxic to biological systems.  
CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).  
CC -1- COFACTOR: Manganese.  
CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE  
CC FAMILY.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL; U11550; AAB08770.1; -;

DR HSP; P17670; IIDS.  
DR InterPro; IPR001189; SODismutase.  
DR Pfam; PF00081; sode; 1.  
DR Pfam; PF02777; sode C; 1.  
DR ProDom; PD000475; SODismutase; 1.  
DR PROSITE; PS00088; SOD\_MN; 1.  
DR Oxidoreductase; Manganese.

KW INIT MET 0  
FT METAL 27 27 BY SIMILARITY.  
FT METAL 75 75 MANGANESE (BY SIMILARITY).  
FT METAL 159 159 MANGANESE (BY SIMILARITY).  
FT METAL 163 163 MANGANESE (BY SIMILARITY).  
SQ SEQUENCE 206 AA; 22912 MW; D6B49083AEBF3C98 CRC64;

Query Match 26.7%; Score 8; DB 1; Length 206;  
Best Local Similarity 100.0%; Pred. No. 0.14;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 ISQINEI 25  
Db 19 ISQINEI 26

RESULT 4  
SODM\_MYCLP STANDARD; PRT; 206 AA.  
ID SODM\_MYCLP  
AC O6165;

DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Superoxide dismutase [Mn] (EC 1.15.1.1).  
GN SODA OR SOD.

OS Mycobacterium lepreum.  
CC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;  
CC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
NCBI\_TaxID=64667;

-RN  
X1 SEQUENCE FROM N.A.  
RC STRAIN=Hawaiian;  
RA Nakamura M.;

RL Submitted (SEP-1992) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: Destroys radicals which are normally produced within the  
CC cells and which are toxic to biological systems.  
CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).  
CC -1- COFACTOR: Manganese (By similarity).  
CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE  
CC FAMILY.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; D13288; BAA28850.1; -;  
DR HSP; P17670; IIDS.  
DR InterPro; IPR001189; SODismutase.  
DR Pfam; PF00081; sode; 1.  
DR Pfam; PF02777; sode C; 1.  
DR ProDom; PD000475; SODismutase; 1.  
DR PROSITE; PS00088; SOD\_MN; 1.  
DR Oxidoreductase; Manganese.

KW INIT MET 0  
FT METAL 27 27 BY SIMILARITY.  
FT METAL 75 75 MANGANESE (BY SIMILARITY).  
FT METAL 159 159 MANGANESE (BY SIMILARITY).  
FT METAL 163 163 MANGANESE (BY SIMILARITY).  
SQ SEQUENCE 206 AA; 23031 MW; 92A063471FA9F22F CRC64;

Query Match 26.7%; Score 8; DB 1; Length 206;  
Best Local Similarity 100.0%; Pred. No. 0.14;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 18 ISGQINEI 25  
 |||||  
 Db 19 ISGQINEI 26

## RESULT 5

SODM\_MYCFO STANDARD; PRT; 206 AA.

AC Q59519; (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Superoxide dismutase [Mn] (EC 1.15.1.1).  
 GN SODA OR SOD.  
 OS Mycobacterium fortuitum.  
 OC Bacteria; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.

NCBI\_TaxID=1766;

[1] SEQUENCE FROM N.A.

RA STRAIN=ATCC 5841;  
 RA MEDLINE=96102872; PubMed=8586279;  
 RA Mendez M.C., Domenech P., Prieto J., Garcia M.J.;

RT "Cloning and expression of the Mycobacterium fortuitum superoxide  
 dismutase gene.";

RL FEMS Microbiol. Lett. 134:273-278(1995).

CC -1- FUNCTION: Destroys radicals which are normally produced within the  
 cells and which are toxic to biological systems.

CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).

CC -1- COFACTOR: Manganese.

CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE  
 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL: X70914; CAA50266.1; -.

DR HSSP: P17670; 1IDS.

DR InterPro: IPR001189; SODismutase.

DR Pfam: PF00081; sodfe\_1.

DR Pfam: PF02777; sodfe\_C; 1.

DR ProDom: PD000475; SODismutase; 1.

DR PROSITE: PS00088; SOD\_MN; 1.

DR Oxidoreductase; Manganese.

FT INIT MET 0

FT METAL 27

FT METAL 75

FT METAL 159

FT METAL 163

SO SEQUENCE 206 AA; 22833 MW; 500625A8B9321246 CRC64;

Query Match 23.3%; Score 7; DB 1; Length 206;  
 Best Local Similarity 100.0%; Pred. No. 1.5;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 18 ISGQINEI 24  
 |||||  
 Db 19 ISGQINEI 25

## RESULT 6

SODM\_MYCSM STANDARD; PRT; 206 AA.

AC P53649; Q8S612;

DT 01-OCT-1996 (Rel. 34, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Superoxide dismutase [Mn] (EC 1.15.1.1).  
 GN SODA OR SOD.  
 OS Mycobacterium smegmatis.  
 OC Bacteria; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.

NCBI\_TaxID=1772;

[1] SEQUENCE FROM N.A., AND CHARACTERIZATION.

RP STRAIN=1-2C;

RC MEDLINE=99134360; PubMed=9933629;

RA Harth G., Horwitz M.A.;

RT "Export of recombinant Mycobacterium tuberculosis superoxide dismutase  
 is dependent upon both information in the protein and mycobacterial  
 export machinery. A model for studying export of leaderless proteins  
 by pathogenic mycobacteria.";

RT J. Biol. Chem. 274:4281-4292(1999).

RL [2] REVISIONS TO 116 AND 201.

RP Tullius M.V., Harth G., Horwitz M.A.;

RA submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

RL [3] SEQUENCE OF 27-164 FROM N.A.

RP STRAIN=NCTC 10265 / ATCC 14468 / W-113;

RC Bull T.J., Shannon D.C., Archard L.C.;

RA "Rapid identification of mycobacteria from AIDS patients by capillary  
 electrophoretic profiling of amplified SOD gene.";

RT J. Clin. Pathol. Clin. Mol. Pathol. 48:124-133(1995).

RL -1- FUNCTION: Destroys radicals which are normally produced within the  
 cells and which are toxic to biological systems.

CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).

CC -1- COFACTOR: Manganese.

CC -1- SUBUNIT: HOMOTETRAMER.

CC -1- SUBCELLULAR LOCATION: PARTIALLY SECRETED.

CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE  
 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL: AF061031; AAD15825.2; -.

DR EMBL: Z48214; CAA88247.1; -.

DR HSSP: P17670; 1IDS.

DR InterPro: IPR001189; SODismutase.

DR Pfam: PF00081; sodfe\_1.

DR Pfam: PF02777; sodfe\_C; 1.

DR ProDom: PD000475; SODismutase; 1.

DR PROSITE: PS00088; SOD\_MN; 1.

DR Oxidoreductase; Manganese.

FT INIT MET 0

FT METAL 27

FT METAL 75

FT METAL 159

FT METAL 163

SO SEQUENCE 206 AA; 22804 MW; B997A1AD0374AEC9 CRC64;

Query Match 23.3%; Score 7; DB 1; Length 206;  
 Best Local Similarity 100.0%; Pred. No. 1.5;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 18 ISGQINEI 24  
 |||||  
 Db 19 ISGQINEI 25

Query Match 23.3%; Score 7; DB 1; Length 206;  
 Best Local Similarity 100.0%; Pred. No. 1.5;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## RESULT 7

SODM\_NOCAS STANDARD; PRT; 206 AA.

AC P53649; Q8S612;

DT 01-OCT-1996 (Rel. 34, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

```

AC P53651;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Superoxide dismutase [Mn] (EC 1.15.1.1).
GN SODA OR SOD.
OS Nocardia asteroides.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Nocardiaceae; Nocardia.
OX NCBI_TaxID=1824;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GUH2;
RX MEDLINE=96060854; PubMed=7590304;
RA Aicendor D.J., Chapman G.D., Beaman B.L.;
RT "Isolation, sequencing and expression of the superoxide dismutase-
encoding gene (sod) of Nocardia asteroides strain GUH-2.";
RL Gene 164143-147(1995).
CC -1- FUNCTION: Destroys radicals which are normally produced within the
cells and which are toxic to biological systems.
CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -1- COFACTOR: Manganese (By similarity).
CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE
FAMILY.
CC -----
CC THE SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U02341; AAA91964.1; -
DR HSSP: P17670; 1IDS.
DR InterPro: IPR001189; SODismutase.
DR Pfam: PF00081; sode; 1.
DR Pfam: PF02777; sode; 1.
DR ProDom: PD000475; SODismutase; 1.
DR PROSITE: PS00088; SOD_MN; 1.
DR Oxidoreductase; Manganese.
KM INIT_MET 0
FT METAL 27 27 BY SIMILARITY.
FT METAL 75 75 MANGANESE (BY SIMILARITY).
FT METAL 159 159 MANGANESE (BY SIMILARITY).
FT METAL 163 163 MANGANESE (BY SIMILARITY).
FT SEQUENCE 206 AA; 22823 MW; FA2BCPF27BB22CBDC CRC64;
Query Match 23.3%; Score 7; DB 1; Length 206;
- Best Local Similarity 100.0%; Pred.No.1.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 ISGQINE 24
DB 19 ISGQINE 25

```

## RESULT 8

```

SODP_MYCTU
ID SODP_MYCTU STANDARD; PRT; 207 AA.
AC P17670; P96231;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Superoxide dismutase [Fe] (EC 1.15.1.1).
GN SODP OR SODA OR SOD OR RV846 OR MT3960 OR MTCY01A6.22C.
OS Mycobacterium tuberculosis, and
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773, 1765;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC SPECIES=M.tuberculosis; STRAIN=Erdmann;
RX MEDLINE=91251768; PubMed=1904126;
RA Zhang Y.;
RT "Genetic analysis of superoxide dismutase, the 23 kilodalton antigen
RT of Mycobacterium tuberculosis.";
RL Mol. Microbiol. 5:381-391(1991).
RN [2]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC SPECIES=M.tuberculosis; STRAIN=Erdmann;
RX MEDLINE=99134360; PubMed=9933629;
RA Harth G., Horwitz M.A.;
RT "Export of recombinant Mycobacterium tuberculosis superoxide dismutase
RT is dependent upon both information in the protein and mycobacterial
RT export machinery. A model for studying export of leaderless proteins
RT by pathogenic mycobacteria.";
RL J. Biol. Chem. 274:4281-4292(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=M.tuberculosis; STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holtroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne K., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=M.tuberculosis; STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwin M.L., Holt D., Hickey E.,
RA Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC SPECIES=M.bovis; STRAIN=BCG / Pasteur;
RA Kimble E., Sanderson R.J., Gill R.E.;
RT "Superoxide dismutase of M. bovis BCG.";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [6]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RC SPECIES=M.tuberculosis;
RX MEDLINE=95182461; PubMed=7877174;
RA Cooper J.B., McIntyre K., Badasso M.O., Wood S.P., Zhang Y.,
RA Garbe T.R., Young D.;
RT "X-ray structure analysis of the iron-dependent superoxide dismutase
RT from Mycobacterium tuberculosis at 2.0-A resolution reveals novel
RT dimer-dimer interactions.";
RL J. Mol. Biol. 246:531-544(1995).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS) OF MUTANT ALA-152.
RC SPECIES=M.tuberculosis;
RX MEDLINE=96244503; PubMed=8674528;
RA Cooper J.B., Seward S., Erskine P.T., Badasso M.O., Wood S.P.,
RA Zhang Y., Young D.;
RT "X-ray structure analysis of an engineered Fe-superoxide dismutase
RT Gly-Ala mutant with significantly reduced stability to denaturation.";
RL FEBS Lett. 387:105-108(1996).
CC -1- FUNCTION: Destroys radicals which are normally produced within the
cells and which are toxic to biological systems.
CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: Secreted.

```



```
CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE
CC FAMILY.
CC CAUTION: ALTHOUGH FOUND EXTRACELLULARLY, NO SIGNAL SEQUENCE IS
CC PRESENT. AN ALTERNATIVE SECRETORY PATHWAY MAY BE USED.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X52861; CAA37042.1; -
DR EMBL; AF061030; AAD15824.1; -
DR EMBL; Z83864; CAB06220.1; -
DR EMBL; AE007188; AAK48327.1; -
DR EMBL; AF077406; AAC27527.1; -
DR PIR; S10908; S10908.
DR PIR; S15205; S15205.
DR PDB; 1IDS; 20-DEC-94.
DR TIGR; MT3960; -
DR Tuberculin; Rv3946; -
DR InterPro; IPR001189; SODismutase.
DR Pfam; PF00081; sode; 1.
DR Pfam; PF02777; sode; 1.
DR ProDom; PD000475; SODismutase; 1.
DR PROSITE; PS00089; SOD_MN; 1.
DR Oxidoreductase; Iron_3D-structure; Complete proteome.
FT METAL 28 28 IRON.
FT METAL 76 76 IRON.
FT METAL 160 160 IRON.
FT METAL 164 164 IRON.
SQ SEQUENCE 207 AA; 23034 MW; DE8F5921DABE54A CRC64;

Query Match 23.3%; Score 7; DB 1; Length 207;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 ISGQINE 24
Db 20 ISGQINE 26

RESULT 9
GCS2_BACSU
ID_GCS2_BACSU STANDARD; PRT; 488 AA.
AC PS4377;
RT 01-OCT-1996 (Rel. 34, Created)
RA 01-OCT-1996 (Rel. 34, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable glycine dehydrogenase [decarboxylating] subunit 2
DE (EC 1.4.4.2) (Glycine decarboxylase) (Glycine cleavage system P-
DE protein).
GN YOHK.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / JH642;
RX MEDLINE=97124195; PubMed=8969508;
RA Mizuno M., Mesuda S., Takemaru K.-I., Hosono S., Sato T., Takeuchi M.,
RA Kobayashi Y.;
RT "Systematic sequencing of the 283 kb 210 degrees-232 degrees region of
RT the Bacillus subtilis genome containing the skin element and many
RT sporulation genes."
RL Microbiology 142:3103-3111(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogatawara N., Moszer I., Albertini A.M., Alloni G.,
```

```
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriss R., Boursier L., Brans A., Braun M., Briqnell S.C., Bron S.,
RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codant J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Enlian K.D., Erlington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Funa S., Gallizi A., Galleron N.,
RA Gilm S.Y., Glaeser P., Goffeau A., Goldthly E.J., Grandi G.,
RA Giuseppe G., Guy B.J., Haga K., Halesch J., Harwood C.R., Henaut A.,
RA Hilbert H., Holappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Kamata D., Kasahara Y., Klier-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kunita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Patro V., Pohl T.M., Portetelle D., Portwollik S., Prescott A.M.,
RA Prescan E., Pujic P., Punnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Serron S.J., Serron P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tanakoshi A., Tanaka T., Terpetre P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vamler F., Vassart A.,
RA Viari A., Wambut R., Wedler E., Wedler H., Weitzneger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yaumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumestein E., Yoshikawa H., Dandini A.,
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis."
RT Nature 350:249-256(1997).
RL Nature 350:249-256(1997).
CC -1- FUNCTION: THE GLYCINE CLEAVAGE SYSTEM CATALYZES THE DEGRADATION OF
CC GLYCINE. THE P PROTEIN BINDS THE ALPHA-AMINO GROUP OF GLYCINE
CC THROUGH ITS PYRIDOXAL PHOSPHATE COFACTOR. CO(2) IS RELEASED AND
CC THE REMAINING METHYLAMINE MOIETY IS THEN TRANSFERRED TO THE
CC LIPONIDE COFACTOR OF THE H PROTEIN (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: Glycine + lipoylprotein = S-
CC -1- cofactor: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC -1- SUBUNIT: THE GLYCINE CLEAVAGE SYSTEM IS COMPOSED OF FOUR PROTEINS:
CC P, T, L, AND H.
CC -1- SIMILARITY: TO THE C-TERMINAL OF OTHER GLYCINE CLEAVAGE SYSTEM P-
CC PROTEIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D84432; BAA12548.1; -
DR EMBL; Z99116; CAB14386.1; -
DR Subtilisin; Bg1511; yqkK.
DR InterPro; IPR003437; GDC-P.
DR Pfam; PF02347; GDC-P; 1.
KW Hypothetical protein; Oxidoreductase; Pyridoxal phosphate;
KW Complete proteome.
FT BINDING 273 273 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 488 AA; 54427 MW; 0B9C7DAD75E657C CRC64;

Query Match 23.3%; Score 7; DB 1; Length 488;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YSLPDL 8
Db 20 YSLPDL 26

RESULT 10
CAPP_SYNY3
ID_CAPP_SYNY3 STANDARD; PRT; 1034 AA.
```





```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X95384; CNA64670.1; -
DR EMBL; AY026764; AAK01939.1; -
DR EMBL; BC008418; AAH08418.1; -
DR EMBL; BC010280; AAH0280.1; -
DR EMBL; BC012592; AAH12592.1; -
DR HSSP; P39330; 1009.
DR MIM; 602487; -
DR InterPro; IPR000543; YjGF-11ke.
DR Pfam; PF01042; UPF0076.1.
DR TIGRPFAMs; TIGR00004; YER05C_YjGF_UK114; 1.
DR PROSITE; PS01094; UPF0076; 1.
KW Nuclear protein.
SQ SEQUENCE 137 AA; 14494 MW; DD0740621E8B6AD CRC64;

Query Match          20.0%; Score 6; DB 1; Length 137;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 YISGQI 22
    |||||
Db 32 YISGQI 37

RESULT 13
SODM_PROFR          STANDARD;      PRT;    201 AA.
ID_P80293;
AC 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Superoxide dismutase [Mn/Fe] (EC 1.15.1.1).
GN SODA.
OS Propionibacterium freudenreichii shermanii.
OC Bacteriia; Actinobacteriia; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Propionibacteriaceae; Propionibacteriaceae;
OC Propionibacterium.
OC NCBI_TaxID=1752;
OX 1;
RN SEQUENCE.
RP STRAIN=PE23;
RC MEDLINE=94139724; PubMed=8307013;
RA Meier B., Sehn A.P., Schinina M.E., Barra D.;
RA "In vivo incorporation of copper into the iron-exchangeable and
RA manganese-exchangeable superoxide dismutase from Propionibacterium
RA shermanii. Amino acid sequence and identity of the protein
RA moieties.";
RT Eur. J. Biochem. 219:463-468 (1994).
RL 4;
RN 12;
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RA Schmidt M., Meier B., Parak F.;
RT "X-ray structure of the cambialistic superoxide dismutase from
RT Propionibacterium shermanii active with Fe or Mn.";
RL J. Biol. Inorg. Chem. 1:533-541 (1996).
RN 13;
RP X-RAY CRYSTALLOGRAPHY (1.55 ANGSTROMS).
RA Schmidt M., Scherk C., Iakovleva O., Nolting H.F., Meier B., Parak F.;
RL Submitted (SSP-1997) to the PDB data bank.
RN 14;
RP X-RAY CRYSTALLOGRAPHY (1.35 ANGSTROMS).
RX MEDLINE=99248073; PubMed=10231372;
RA Schmidt M.;
RT "Manipulating the coordination number of the ferric iron within the
RT cambialistic superoxide dismutase of Propionibacterium shermanii by
RT changing the pH-value. A crystallographic analysis.";
RL Eur. J. Biochem. 262:117-127 (1999).
CC -1- FUNCTION: Destroys radicals which are normally produced within the

```

```

CC cells and which are toxic to biological systems.
CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -1- COFACTOR: ACTIVE WITH EITHER MANGANESE OR IRON.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE
CC FAMILY.
DR PIR; S41106; S41106.
DR PDB; 1AR4; 12-NOV-97.
DR PDB; 1AR5; 12-NOV-97.
DR PDB; 1AVM; 18-MAR-98.
DR PDB; 1BS3; 15-JUN-99.
DR PDB; 1BSW; 15-JUN-99.
DR PDB; 1BT8; 15-JUN-99.
DR InterPro; IPR001189; SODismutase.
DR Pfam; PF00081; sode; 1.
DR Pfam; PF02777; sodeC; 1.
DR ProDom; PD000475; SODismutase; 1.
DR PROSITE; PS00086; SOD_MN; 1.
KW Oxidoreductase; Manganese; Iron; 3D-structure.
FT METAL 27 27 MANGANESE OR IRON.
FT METAL 75 75 MANGANESE OR IRON.
FT METAL 161 161 MANGANESE OR IRON.
FT METAL 165 165 MANGANESE OR IRON.
SQ SEQUENCE 201 AA; 22633 MW; 5BF8F424C7B32E00 CRC64;

Query Match          20.0%; Score 6; DB 1; Length 201;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 EPIYSG 20
    |||||
Db 16 EPIYSG 21

RESULT 14
3MGH_LISMO          STANDARD;      PRT;    207 AA.
ID_3MGH_LISMO
AC P58621;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative 3-methyladenine DNA glycosylase (EC 3.2.2.-).
GN LMO00928.
OS Listeria monocytogenes.
OC Bacteriia; Firmicutes; Bacillales; Listeriaceae; Listeria.
OC NCBI_TaxID=1639;
OX 1;
RN SEQUENCE FROM N.A.
RP STRAIN=EGD-e / Serovar 1/2a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaeser P., Frangeul L., Buchrieser C., Ruenick C., Amend A.;
RA Baquero F., Berche P., Bloecher H., Brandt P., Chakraborty T.,
RA Charbit A., Cherouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Duran L., Duesetget O.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapat G.,
RA Medueno E., Maitournan A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordsiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Weiland J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852 (2001).
CC -1- SIMILARITY: BELONGS TO THE DNA GLYCOSYLASE MFG FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

DR EMBL, AL591977; CAC99006.1; -.  
 DR L18L1818; LMO00928; -.  
 DR InterPro: IPR003180; Pair DNA\_glyco.  
 DR TIGRPFAMS; TIGR00567; 3mg; 1.  
 KM Hypothetical protein; DNA repair; Hydrolyase; Complete proteome.  
 SQ SEQUENCE 207 AA; 23372 MW; 8066CCF4A5E1807D CRC64;

Query Match 20.0%; Score 6; DB 1; Length 207;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 PYISQ 21  
 DB 191 PYISQ 196

RESULT 15  
 SODM\_CHLPHN STANDARD; PRT; 207 AA.  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Superoxide dismutase [Mn] (EC 1.15.1.1).  
 GN SODA OR SODM OR CPN0057 OR CP0718.  
 OS Chlamydia pneumoniae (Chlamydiophila pneumoniae).  
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
 OX NCBI\_TaxID=83558;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=EWL029;  
 RX MEDLINE=9206606; PubMed=10192388;  
 RA Kilmann S., Mitchell W.J., Marathe C., Fan J., Hyman R.W.,  
 RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;  
 RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis";  
 RN Nat. Genet. 21:385-389(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AR39;  
 RX MEDLINE=20150255; PubMed=10684935;  
 RA Read T.D., Brumham R.C., Shen C., Gill S.R., Heidelberg J.F.,  
 RA White O., Hickey E.K., Peterson J., Uettersack T., Berry K., Baes S.,  
 RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,  
 RA Gilm M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,  
 RA Eisen J., Fraser C.M.;  
 RT "Genome sequences of Chlamydia trachomatis Mohn and Chlamydia  
 pneumoniae AR39";  
 RN Nucleic Acids Res. 28:1397-1406(2000).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=J138;  
 RX MEDLINE=20330349; PubMed=10871362;  
 RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,  
 RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;  
 RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138  
 from Japan and CML029 from USA";  
 RN Nucleic Acids Res. 28:2311-2314(2000).  
 RL -!- FUNCTION: Destroys radicals which are normally produced within the  
 cells and which are toxic to biological systems.  
 CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).  
 CC -!- COFACTOR: Manganese (By similarity).  
 CC -!- SUBUNIT: HOMODIMER (By similarity).  
 CC -!- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE  
 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).

DR EMBL, AE001591; AAD18210.1; -.  
 DR EMBL, AE002230; AAF38524.1; -.  
 DR EMBL, AP002545; BAA98268.1; -.  
 DR HSSP; P04179; 1AP6.  
 DR PHCI-2DPAGE; Q929C4; -.  
 DR TIGR, CP0718; -.  
 DR InterPro: IPR001189; SODismutase.  
 DR Pfam; PF00081; sode; 1.  
 DR Pfam; PF02777; sode; C; 1.  
 DR Prodom; PD000475; SODismutase; 1.  
 DR PROSITE; PS00088; SOD MN; 1.  
 KM Oxidoreductase; Manganese; Complete proteome.  
 FT METAL 31 31 MANGANESE (BY SIMILARITY).  
 FT METAL 78 78 MANGANESE (BY SIMILARITY).  
 FT METAL 166 166 MANGANESE (BY SIMILARITY).  
 FT METAL 170 170 MANGANESE (BY SIMILARITY).  
 SQ SEQUENCE 207 AA; 23541 MW; 45A4A6FF096F8934 CRC64;

Query Match 20.0%; Score 6; DB 1; Length 207;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YSLPEL 7  
 DB 6 YSLPEL 11

RESULT 16  
 PURU\_HAEIN STANDARD; PRT; 278 AA.  
 ID PURU\_HAEIN  
 AC Q03432;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Formyltetrahydrofolate deformylase (EC 3.5.1.10) (Formyl-FH(4)  
 DE hydrolase).  
 GN PURU OR H11588.  
 OS Haemophilus influenzae.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
 OC Haemophilus.  
 OX NCBI\_TaxID=727;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Rd / KW20 / ATCC 51907;  
 RX MEDLINE=95350630; PubMed=7542800;  
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,  
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
 RA Uettersack T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,  
 RA Fine L.D., Frichman J.L., Fuhrmann J.L., Geoghegan N.S.M.,  
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
 RA Venter J.C.;  
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae  
 Rd.";  
 RN Science 269:496-512(1995).  
 RN [2]  
 RP SEQUENCE OF 64-278 FROM N.A.  
 RC STRAIN=RM 7004 / Serotype B;  
 RX MEDLINE=93328119; PubMed=8335255;  
 RA Maskell D.J.;  
 RT "Cloning and sequencing of the Haemophilus influenzae ara gene";  
 RN Gene 129:155-156(1993).  
 RL -!- FUNCTION: PRODUCES FORMATE FROM FORMYL-TETRAHYDROFOLATE. PROVIDES  
 CC THE MAJOR SOURCE OF FORMATE FOR THE PURT-DEPENDENT SYNTHESIS OF  
 CC 5'-PHOSPHORIBOSYL-N-FORMYLGLYCINAMIDE (FGAR) DURING AEROBIC  
 CC GROWTH. HAS A ROLE IN REGULATING THE ONE-CARBON POOL  
 CC (BY SIMILARITY).  
 CC -!- CATALYTIC ACTIVITY: 10-formyltetrahydrofolate + H(2)O = formate +  
 CC tetrahydrofolate.  
 CC -!- ENZYME REGULATION: ACTIVATED BY METHIONINE, INHIBITED BY GLYCINE  
 CC (BY SIMILARITY).

```

CC -1- PATHWAY: De novo purine biosynthesis.
CC -1- SUBUNIT: HOMOMEXAMER (BY SIMILARITY).
CC -1- SIMILARITY: SOME, TO GAT TRANSFORMYLASE (PURN).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U32833; AAC23236.1; -.
CC EMBL; L04686; AAA24942.1; -.
CC PIR; P06066; P06066.
CC HSP; P08179; IGRC.
CC TIGR; H15688; -.
CC InterPro; IPR002912; ACT.
CC InterPro; IPR004810; PURU.
CC InterPro; IPR002376; formyl transf.
CC Pfam; PF00551; formyl_transf; 1.
CC ORFam; PF01842; ACT; 1.
CC TIGRfam; TIGR00655; PurU; 1.
CC Purine biosynthesis; Hydrolyase; One-carbon metabolism;
CC Complete proteome.
CC ACT_SITE 223 BY SIMILARITY.
CC CONFLICT 115 VIG -> RNR (IN REF. 2).
CC CONFLICT 140 HEN -> PK (IN REF. 2).
CC CONFLICT 205 K -> E (IN REF. 2).
CC SEQUENCE 278 AA; 32173 MW; 7F375AB3C4225C4B CRC64;

Query Match 20.0%; Score 6; DB 1; Length 278;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KXSLEP 6
Db 66 KXSLEP 71

RESULT 17
MENC_MYCLE STANDARD; PRT; 334 AA.
AC Q9CBB2;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Probable O-succinylbenzoate-CoA synthase (EC 4.2.1.-) (OSB synthase)
DE (OSBS) (4-(2-carboxyphenyl)-4-oxobutyric acid synthase) (O-
DE succinylbenzoyl-CoA synthase).
DE MENC OR ML2268.
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1169;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TN;
RX MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Eiglmeyer K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garner T., Churchill T., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltham T., Fraser A., Hamlin N.,
RA Holtrop S., Hornby T., Jegerle K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrett B.G.;
RT "Massive gene decay in the leprosy bacillus.";
RL Nature 409:1007-1011 (2001).
CC -1- FUNCTION: Converts SHHC to OSB (By similarity).
CC -1- CATALYTIC ACTIVITY: 2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-
CC carboxylate + O-succinylbenzoate + H(2)O.

```

```

CC -1- COFACTOR: Requires a divalent metal ion (By similarity).
CC -1- PATHWAY: Menaguinone biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE MANDULATE RACEMASE / MUONATE
CC LACTONIZING ENZYME FAMILY. MENC SUBFAMILY 1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AL583925; CAC31784.1; -.
CC EMBL; ML2268; -.
CC InterPro; IPR001354; MR_MLE.
CC Pfam; PF01188; MR_MLE; 1.
CC Menaguinone biosynthesis; Lyase; Complete proteome.
CC ACT_SITE 107 BY SIMILARITY.
CC ACT_SITE 109 BY SIMILARITY.
CC SEQUENCE 334 AA; 35310 MW; 9590196DACF93145 CRC64;

Query Match 20.0%; Score 6; DB 1; Length 334;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 LPELDY 9
Db 252 LPELDY 257

RESULT 18
UTR2_YEAST STANDARD; PRT; 347 AA.
AC P32623;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE UTR2 protein (Unknown transcript 2 protein).
DE UTR2 OR YEL040W OR SYGP-ORF18.
GN Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Mulligan J.T., Dietrich F.S., Hennessey K.M., Sehl P., Komp C.,
RA Wei Y., Taylor P., Nakahara K., Roberts D., Davis R.W.;
RL Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.
RX [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,
RA Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,
RA Chung E., Duncan M., Guzman E., Hartzell G., Hunnicke-Smith S.,
RA Hyman R., Kayser A., Komp C., Lashkari D., Lew H., Lin D.,
RA Moseedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,
RA Petrel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,
RA Taylor P., Wei Y., Yelton M., Botstein D., Davis R.W.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
RX [3]
RP SEQUENCE FROM N.A.
RC STRAIN=B-6441;
RX MEDLINE=94016558; PubMed=8411151;
RA Melnick L., Sherman F.;
RT "The gene clusters AKC and COR on chromosomes 5 and 10, respectively,
RT of Saccharomyces cerevisiae share a common ancestry.";
RL J. Mol. Biol. 233:372-381 (1993).
CC -1- SIMILARITY: SOME, TO YEAST YGR189C.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U18779; AAB65002.1; ALT_INT.
DR EMBL; L22173; AAA34941.1; -.
DR EMBL; S65964; AAD13975.1; -.
DR EMBL; S66130; AAB28444.1; -.
DR PIR; S30839; S30839.
DR HSP; P23904; IAO.
DR SGD; S0000766; UTR2.
DR InterPro; IPR00757; Glyco_hydro_16.
DR Pfam; PF00722; Glyco_hydro_16; 1.
FT DOMAIN 234 322 SER-RICH.
FT DOMAIN 269 283 POLY-SER.
FT CONFLICT 10 10 L -> V (IN REF. 3).
FT CONFLICT 171 171 A -> R (IN REF. 3).
FT CONFLICT 234 234 S -> C (IN REF. 3).
SEQUENCE 347 AA; 36692 MW; 1B8AFB862C4BB328 CRC64;

Query Match 20.0%; Score 6; DB 1; Length 347;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ELDYEF 11
    |||||
Db 46 ELDYEF 51

RESULT 19
ID12 HALN1 STANDARD; PRT; 360 AA.
AC OSHHE4; O54623;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE (Isopentenyl-diphosphate delta-isomerase) (EC 5.3.3.2) (IPP isomerase)
DE (Isopentenyl pyrophosphate isomerase).
GN (FNI1 OR VNG5084G OR H0660) AND (FNI2 OR VNG5213G OR H1696) AND
GN (FNI3 OR VNG6081G) AND (FNI4 OR VNG6445G).
OS Halobacterium sp. (Strain NRC-1).
OC Archaea; Euryarchaeota; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_Taxid=64091;
[1]
SEQUENCE FROM N.A.
PLASMID=PNRC100;
RA NG W.V., Clute S.A., Smith T.M., Bungarner R.E., Baskin D., Faust J.,
RA Hall B., Loretz C., Sero J., Slagel J., Hood L., Dassarma S.;
RT "Snapshot of a large dynamic replicon in a halophilic archaeon:
RT megaplasmid or minichromosome?";
RL Genome Res. 8:1131-1141(1998).
RN [2]
SEQUENCE FROM N.A.
RC PLASMID=PNRC200;
RA NG W.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,
RA Shukla H.D., Laaky S.R., Balliga N.S., Thorsson V., Sproga J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Weir R., Goo Y.A.,
RA Leitner B., Keller K., Cruz R., Danon M.J., Hough D.W.,
RA Maddocks D.G., Jabloncki P.E., Krebs M.P., Angewine C.M., Dale H.,
RA Isenberger T.A., Peck R.F., Fehlechner M., Spudich J.L., Jung K.-H.,
RA Alam M., Freltas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RT "Genome sequence of Halobacterium species NRC-1."
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
CC -1- FUNCTION: CATALYZES THE 1,3-ALLYLIC REARRANGEMENT OF THE
CC HOMOMALLYLIC SUBSTRATE ISOPENTENYL (IPP) TO ITS ALLYLIC ISOMER,
CC DIMETHYLLALLYL DIPHOSPHATE (DMAPP) (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: Isopentenyl diphosphate = dimethylallyl
```

```
CC diphosphate.
CC -1- COFACTOR: FMN AND NADPH (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE IPP ISOMERASE TYPE 2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF016485; AAC82844.1; ALT_INT.
DR EMBL; AF016485; AAC82933.1; ALT_INT.
DR EMBL; AE005145; AAG20768.1; -.
DR EMBL; AE005169; AAG21040.1; -.
DR InterPro; IPR003009; FMN_enzyme.
DR InterPro; IPR000262; FMN_hydryac_dh.
DR Pfam; PF01070; FMN_dh; 1.
KM Isomerase; Isoprene biosynthesis; Flavoprotein; FMN; NADP; Plasmid;
SO Complete proteome.
SEQUENCE 360 AA; 38342 MW; 67B14178071D3164 CRC64;

Query Match 20.0%; Score 6; DB 1; Length 360;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 LPELDY 9
    |||||
Db 42 LPELDY 47

RESULT 20
HEMY HAEMIN STANDARD; PRT; 428 AA.
ID HEMY HAEMIN STANDARD; PRT; 428 AA.
AC P44772;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hemy protein homolog.
DE HEMY OR H10602.
GN HEMY OR H10602.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_Taxid=727;
[1]
SEQUENCE FROM N.A.
RC STRAIN=Rd / KM20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kervatage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd."
RL Science 269:496-512(1995).
CC -1- FUNCTION: INVOLVED IN A LATE STEP OF PROTOHEME IX SYNTHESIS
CC (BY SIMILARITY).
CC -1- PATHWAY: Porphyrin biosynthesis.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC -----
DR EMBL; U32742; AAC22260.1; -.
DR TIGR; H10602; -.
DR InterPro; IPR005254; Heme_bio.
DR TIGRPRAS; TIGR00540; hemy_coli; 1.
DR Porphylin biosynthesis; Complete proteome.
SQ SEQUENCE 428 AA; 48776 MW; 8654FAC94689CB5C CRC64;

Query Match 20.0%; Score 6; DB 1; Length 428;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 PYISQ 21
Db 19 PYISQ 24

RESULT 21
TC21_AVEA STANDARD; PRT; 535 AA.
ID TC21_AVEA
PA0412;
RC 01-FEB-1995 (Rel. 31, Created)
RX MEDLINE=94085629; PubMed=7903257;
RA Ethman B., Krenz W., Mummert E., Schaefer E.;
RT "Two TcP-1-related but highly divergent gene families exist in oat
  encoding proteins of assumed chaperone function.";
RL FEBS Lett. 336:313-316(1993).
CC -1- FUNCTION: MOLECULAR CHAPERONE; ASSIST THE FOLDING OF PROTEINS UPON
  ATP HYDROLYSIS. KNOWN TO PLAY A ROLE, IN VITRO, IN THE FOLDING OF
  ACTIN AND TUBULIN.
CC -1- SUBUNIT: HETERO-OLIGOMERIC COMPLEX OF ABOUT 850 TO 900 kDa THAT
  FORMS TWO STACKED RINGS, 12 TO 16 NM IN DIAMETER.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE TCP-1 CHAPERONIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X75777; CAAS3396.1; -.
DR PIR; S39319; S39319.
DR PIR; S40461; S40461.
DR HSP; P48425; 1A6D.
DR InterPro; IPR002194; Chaperonin_TCP-1.
DR InterPro; IPR002423; Cpn60/TCP-1.
DR Pfam; PF00118; Cpn60 TCP1.1.
DR PRINTS; PR00304; TCOMPLEXTCP1.
DR PROSITE; PS00750; TCP1_1; 1.
DR PROSITE; PS00751; TCP1_2; 1.
DR PROSITE; PS00995; TCP1_3; 1.
DR Chaperone; ATP-binding; Multigene family.
SQ SEQUENCE 535 AA; 58899 MW; 6AFA847CA891BC32 CRC64;

Query Match 20.0%; Score 6; DB 1; Length 535;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 9 YFSAT 14
Db 151 YFSAT 156

RESULT 22
TC22_AVEA STANDARD; PRT; 535 AA.
ID TC22_AVEA
P54411;
RC 01-OCT-1996 (Rel. 34, Created)
RX MEDLINE=94085629; PubMed=7903257;
RA Ethman B., Krenz W., Mummert E., Schaefer E.;
RT "Two TcP-1-related but highly divergent gene families exist in oat
  encoding proteins of assumed chaperone function.";
RL FEBS Lett. 336:313-316(1993).
CC -1- FUNCTION: MOLECULAR CHAPERONE; ASSIST THE FOLDING OF PROTEINS UPON
  ATP HYDROLYSIS. KNOWN TO PLAY A ROLE, IN VITRO, IN THE FOLDING OF
  ACTIN AND TUBULIN.
CC -1- SUBUNIT: HETERO-OLIGOMERIC COMPLEX OF ABOUT 850 TO 900 kDa THAT
  FORMS TWO STACKED RINGS, 12 TO 16 NM IN DIAMETER.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE TCP-1 CHAPERONIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X75778; CAAS3397.1; -.
DR HSP; P48424; 1A6D.
DR InterPro; IPR002194; Chaperonin_TCP-1.
DR InterPro; IPR002423; Cpn60/TCP-1.
DR Pfam; PF00118; Cpn60 TCP1.1.
DR PRINTS; PR00304; TCOMPLEXTCP1.
DR PROSITE; PS00750; TCP1_1; 1.
DR PROSITE; PS00751; TCP1_2; 1.
DR PROSITE; PS00995; TCP1_3; 1.
DR Chaperone; ATP-binding; Multigene family.
SQ SEQUENCE 535 AA; 59034 MW; 02761226F5F17B81 CRC64;

Query Match 20.0%; Score 6; DB 1; Length 535;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 YFSAT 14
Db 151 YFSAT 156

RESULT 23
YDDA_ECOLI STANDARD; PRT; 561 AA.
ID YDDA_ECOLI
P31826; P76133; P76876;
RC 01-JUL-1993 (Rel. 26, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical ABC transporter ATP-binding protein ydda (CDS102).
GN YDDA OR B1496.
```

OC Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / MG1655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 RA Blatner F.R., Plunkett G. II, Bloch C.A., Perna N.T., Burland V.,  
 RA Blythe R., Dialo-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.,  
 RT "The complete genome sequence of Escherichia coli K-12."  
 RL Science 277:1453-1474(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12;  
 RX MEDLINE=97251357; PubMed=9097039;  
 RA Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,  
 RA Itoh T., Kawai H., Kashimoto K., Kimura S., Kitakawa M.,  
 RA Kitagawa M., Makino K., Miki T., Mizobuchi K., Mori H., Mori T.,  
 RA Motomura K., Nakade S., Nakamura Y., Nishimoto H., Nishio Y.,  
 RA Oshima T., Saito N., Sampei G., Seki Y., Sivasubraman S.,  
 RA Yamamoto Y., Horiiuchi T.,  
 RA Yamamoto Y., Takeuchi K., Takeuchi Y., Wada C.,  
 RT "A 570-Kb DNA sequence of the Escherichia coli K-12 genome  
 RT corresponding to the 28.0-40.1 min region on the linkage map."  
 RL DNA Res. 3:363-377(1996).  
 RN [3]  
 RP SEQUENCE OF 306-561 FROM N.A.  
 RC STRAIN=K12;  
 RA Thirlin E., Gasser F., Bienville F.;  
 RT "Sequence and functional analysis of an Escherichia coli DNA fragment  
 RT able to complement pggE and pggF from Methylobacterium organophilum."  
 RL Submitted (May-1993) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.  
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AB00246; AAC74569.1; -;  
 DR EMBL: X71917; CA50732.1; ALT\_INIT.  
 DR EMBL: D90791; BAA15167.1; ALT\_INIT.  
 DR EMBL: D90792; BAA15170.1; -;  
 DR PIR: S33468; S33468.  
 DR Ecogen: EG11742; Ydda.  
 DR InterPro: IPR003439; ABC\_transport.  
 DR Pfam: PF00005; ABC\_tran; 1.  
 DR ProDom: PD000006; ABC\_transport; 1.  
 DR PROSITE: PS00211; ABC\_TRANSPORTER; 1.  
 KM Hypothetical protein; ATP-binding; Transport; Transmembrane;  
 KM Complete proteome.  
 FT TRANSMEM 4 24 POTENTIAL.  
 FT TRANSMEM 32 52 POTENTIAL.  
 FT TRANSMEM 71 91 POTENTIAL.  
 FT TRANSMEM 152 172 POTENTIAL.  
 FT TRANSMEM 188 208 POTENTIAL.  
 FT TRANSMEM 291 311 POTENTIAL.  
 FT NP\_BIND 400 407 ATP (POTENTIAL).  
 SQ SEQUENCE 561 AA; 64984 MW; 92B47426294413P5 CRC64;

Db 302 ISGQIN 307  
 RESULT 24  
 EMBL\_MOUSE  
 ID EMBL\_MOUSE STANDARD; PRT; 879 AA.  
 AC Q922H5;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DE 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Band 4.1-like protein 1 (Neuronal protein 4.1) (4.1N).  
 GN EPB41L1 OR EPB4.1L1 OR EPB4.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=99343811; PubMed=10414974;  
 RA Walensky L.D., Blackshaw S., Liao D., Mackins C.C., Weier H.-U.G.,  
 RA Parra M., Huganir R.L., Conboy J.G., Mohandas N., Snyder S.H.;  
 RT "A novel neuron-enriched homolog of the erythrocyte membrane  
 RL cytoskeletal protein 4.1."  
 RT J. Neurosci. 19:6457-6467(1999).  
 CC -1- FUNCTION: May function to confer stability and plasticity to  
 CC neuronal membrane via multiple interactions, including the  
 CC spectrin-actin-based cytoskeleton. Integral membrane channels and  
 CC membrane-associated guanylate kinases.  
 CC -1- TISSUE SPECIFICITY: Highest expression in brain, also present in  
 CC kidney, olfactory epithelium, retina, sensory ganglia,  
 CC gastrointestinal tract (only enteric neurons) and lung.  
 CC -1- SIMILARITY: CONTAINS 1 BAND 4.1-LIKE DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF061283; AAC68583.1; -;  
 DR MGD: MGI:103010; Epp4.1l1.  
 DR InterPro: IPR000299; Band\_4.1.  
 DR Pfam: PF00373; Band\_41; 1.  
 DR PRINTS: PR00935; BAND41.  
 DR SMART: SM00295; B41; 1.  
 DR PROSITE: PS00660; BAND\_41\_1; 1.  
 DR PROSITE: PS00661; BAND\_41\_2; 1.  
 DR PROSITE: PS00657; BAND\_41\_3; 1.  
 KM Structural protein; Cytoskeleton.  
 FT DOMAIN 94 308 BAND 4.1-LIKE.  
 FT DOMAIN 375 482 HYDROPHILIC.  
 FT DOMAIN 483 541 SPECTRIN-ACTIN-BINDING.  
 FT DOMAIN 743 879 CARBOXYL-TERMINAL (CTD).  
 SQ SEQUENCE 879 AA; 98314 MW; 2B5014B4DCA325 CRC64;

Query Match 20.0%; Score 6; DB 1; Length 879;  
 Best Local Similarity 100.0%; Pred. No. 65;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SLPELD 8  
 Db 639 SLPELD 644

RESULT 25  
 EMBL\_HUMAN  
 ID EMBL\_HUMAN STANDARD; PRT; 881 AA.  
 AC Q9H4G0; O15046;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)



DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Band 4.1-like protein 1 (Neuronal protein 4.1) (4.1N).  
 GN EPB41.1 OR KIAA0338.  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 ON NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain.  
 RX MEDLINE=97349984; PubMed=9205841;  
 RA Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N.,  
 RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.,  
 RT "Prediction of the coding sequences of unidentified human genes. VII.  
 RT The complete sequences of 100 new cDNA clones from brain which can  
 RT code for large proteins in vitro."  
 RL Dna Res. 4:141-150(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21638749; PubMed=11780052;  
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
 RA Jones M., Stevrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,  
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
 RA Chapman J.C., Clamp M., Collier R.E., Connor R.E., Corby N.R.,  
 RA Cragg S., Copley V.E., Collier R.E., Clark L.N., Clark S.Y., Clee C.M.,  
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,  
 RA Ellington A.G., Frankland J.A., Frazer A., French L., Garner P.,  
 RA Griffiths D.V., Griffiths C., Griffiths M.N.D., Gilliam R., Hall R.E.,  
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird D.M., Lawlor S.,  
 RA Levelepain M.H., Leverish M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
 RA Marsh V.L., Martin S.L., McCormack L.J., McElroy K., McMurtry A.A.,  
 RA Oliver S.A., Mistry D., Moore M.J.P., Mullikin J.C., Nickerson T.,  
 RA Oline K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,  
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,  
 RA Skuse C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,  
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
 RA Tracey A., Tromans A.C., Vaudin R., Wall M., Wallis J.M.,  
 RA Whitehead S.L., Whiteaker P., Willey D.L., Williams L., Williams S.A.,  
 RA Wilmshire J., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
 RA Rogers J.;  
 RT "The DNA sequence and comparative analysis of human chromosome 20."  
 RL Nature 414:865-871(2001).  
 CC -1- FUNCTION: May function to confer stability and plasticity to  
 CC neuronal membrane via multiple interactions, including the  
 CC spectrin-actin-based cytoskeleton, integral membrane channels and  
 CC membrane-associated guanylate kinases.  
 CC -1- TISSUE SPECIFICITY: Highest expression in brain, lower in heart,  
 CC kidney, pancreas, placenta, lung and skeletal muscle.  
 CC -1- SIMILARITY: CONTAINS 1 BAND 4.1-LIKE DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AB002336; BAA20796.1; ALT\_INIT.  
 DR EMBL; AL121895; CAC09920.1; --  
 DR GenBank; HGNC:3378; EPB41.1.  
 DR MIM; 602879; --  
 DR InterPro; IPR000299; Band 4.1.  
 DR Pfam; PF00373; Band 41; 1.  
 DR PRINTS; PR00935; BAND41.  
 DR SMART; SM00295; B41; 1.  
 DR PROSITE; PS00660; BAND\_41\_1; 1.  
 DR PROSITE; PS00661; BAND\_41\_2; 1.

DR PROSITE; PSS0057; BAND\_41\_3; 1.  
 KW Structural protein; Cytoskeleton.  
 FT DOMAIN 94 308 BAND 4.1-LIKE.  
 FT DOMAIN 483 541 SPECTRIN-ACTIN-BINDING.  
 FT DOMAIN 746 881 CARBOXYL-TERMINAL (CTD).  
 FT CONFLICT 484 484 MISSING (IN REF. 2).  
 FT CONFLICT 729 729 MISSING (IN REF. 2).  
 SQ SEQUENCE 881 AA; 98502 MW; D923CF554EDB41D3 CRC64;  
 Query Match 20.0%; Score 6; DB 1; Length 881;  
 Best Local Similarity 100.0%; Pred. No. 66;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 639 SLFELD 644  
 Qy 3 SLFELD 8  
 Db 639 SLFELD 644  
 RESULT 26  
 UVRA\_NEIMA STANDARD; PRT; 948 AA.  
 ID UVRA\_NEIMA  
 AC 09JUS4:  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Excinuclease ABC subunit A.  
 GN UVRA OR NMA1159.  
 OS Neisseria meningitidis (serogroup A).  
 CC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.  
 ON NCBI\_TaxID=55699;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=22491 / Serogroup A / Serotype 4A;  
 RX MEDLINE=20222556; PubMed=10761919;  
 RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,  
 RA Klee S.R., Morelli G., Bauman D., Brown D., Chillingworth T.,  
 RA Davies R.M., Davis P., Devlin K., Felwell T., Hamlin N., Holtroyd S.,  
 RA Jorgels K., Leather S., Moule S., Mungall K., Quail M.A.,  
 RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,  
 RA Whitehead S., Spratt B.G., Barrall B.G.;  
 RT "Complete DNA sequence of a serogroup A strain of Neisseria  
 RT meningitidis 22491."  
 RL Nature 404:502-506(2000).  
 CC -1- FUNCTION: The ABC excision nuclease is a DNA repair enzyme that  
 CC catalyzes the excision reaction of UV-damaged nucleotide segments  
 CC producing oligomers having the modified base(s). UvrA is an ATPase  
 CC and a DNA-binding protein that preferentially binds single-  
 CC stranded or UV-irradiated double-stranded DNA (by similarity).  
 CC -1- SUBUNIT: Consists of three subunits; uvrA, uvrB and uvrC.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. UVRA SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AL162755; CAB84421.1; --  
 DR InterPro; IPR003439; ABC transportr.  
 DR InterPro; IPR004602; UvrA.  
 DR Pfam; PF00005; ABC\_tran; 2.  
 DR ProDom; PD000006; ABC\_transportr; 1.  
 DR TIGRFAMs; TIGR00630; uvrA; 1.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER; 2.  
 KW SOS response; Excision nuclease; DNA repair; ATP-binding; Repeat;  
 KW DNA-binding; Zinc-finger; Complete proteome.  
 FT NP\_BIND 42 49 ATP (POTENTIAL).  
 FT NP\_BIND 649 656 ATP (POTENTIAL).  
 FT ZN\_FING 262 289 C4-TYPE.  
 FT ZN\_FING 748 774 C4-TYPE.



SEQ SEQUENCE 948 AA; 105191 MW; E98717DF62B9A0D CRC64;

Query Match 20.0%; Score 6; DB 1; Length 948;

Best Local Similarity 100.0%; Pred. No. 70;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YSLEP 7

Db 267 YSLEP 272

RESULT 27

UVRA\_NEIMB

ID UVRA\_NEIMB STANDARD; PRT; 949 AA.

AC Q9JZP1;

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

EXCINUCLEASE ABC subunit A.

UVRA OR NMB0962.

OS Neisseria meningitidis (serogroup B).

OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.

OX NCBI\_TaxID=491;

[1]

SEQUENCE FROM N.A.

RC STRAIN=MC58 / Serogroup B;

RA MEDLINE=20175755; PubMed=10710307;

RA Tettein H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,

RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,

RA Nelson W.C., Gwin M.L., Deboy R., Peterson J.D., Hickey E.K.,

RA Hatt D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,

RA Mason T., Ciecko A., Parksey D.S., Blair E., Cifton H., Clark E.B.,

RA Cotton M.D., Uterback T.R., Khouri H., Qin H., Vamshayan J.,

RA Gyll J., Scarlato V., Maignani V., Pizza M., Grandi G., Sun L.,

RA Smith H.O., Frazer C.M., Moxon E.R., Rappuoli R., Venter J.C.,

RT "Complete genome sequence of Neisseria meningitidis serogroup B strain

MC58."

RL Science 287:1809-1815 (2000).

CC -1- FUNCTION: The ABC excision nuclease is a DNA repair enzyme that

CC catalyzes the excision reaction of UV-damaged nucleotide segments

CC producing oligomers having the modified base(s). UVRA is an ATPase

CC and a DNA-binding protein that preferentially binds single-

CC stranded or UV-irradiated double-stranded DNA (by similarity).

CC -1- SUBUNIT: Consists of three subunits; uvra, uvrb and uvrc.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. UVRA SUBFAMILY.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

CC

DR EMBL; AE002447; AAF41368.1; -.

DR TIGR; NMB0962; -.

DR InterPro; IPR003439; ABC\_transportr.

DR InterPro; IPR004602; UVRA.

DR Pfam; PF00005; ABC\_tran; 2.

DR ProDom; PD000006; ABC\_transportr; 1.

DR TIGRFAMs; TIGR00630; uvra; 1.

DR PROSITE; PS00211; ABC\_TRANSPORTER; 2.

DR SOS response; Excision nuclease; DNA repair; ATP-binding; Repeat;

KW DNA-binding; Zinc-finger; Complete proteome.

KW NP\_BIND 42 49 ATP (POTENTIAL).

FT NP\_BIND 649 656 ATP (POTENTIAL).

FT ZN\_FING 262 289 C4-TYPE.

FT ZN\_FING 748 774 C4-TYPE.

SO SEQUENCE 949 AA; 105536 MW; 153F5AE2D48CD035 CRC64;

Query Match 20.0%; Score 6; DB 1; Length 949;

Best Local Similarity 100.0%; Pred. No. 70;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YSLEP 7

Db 267 YSLEP 272

RESULT 28

UVRA\_NEIGO

ID UVRA\_NEIGO STANDARD; PRT; 950 AA.

AC O50968;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Excinuclease ABC subunit A.

GN UVRA.

OS Neisseria gonorrhoeae.

OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.

OX NCBI\_TaxID=485;

[1]

SEQUENCE FROM N.A.

RX MEDLINE=97340932; PubMed=9197406;

RA Black C.G., Pyfe J.A.M., Davies J.K.;

RT "Cloning, nucleotide sequence and transcriptional analysis of the

RT uvra gene from Neisseria gonorrhoeae."

RL Mol. Genet. 254:479-485 (1997).

CC -1- FUNCTION: THE ABC EXCISION NUCLEASE IS A DNA REPAIR ENZYME THAT

CC CATALYZES THE EXCISION REACTION OF UV-DAMAGED NUCLEOTIDE SEGMENTS

CC PRODUCING OLIGOMERS HAVING THE MODIFIED BASE(S). UVRA IS AN ATPASE

CC AND A DNA-BINDING PROTEIN THAT PREFERENTIALLY BINDS SINGLE-

CC STRANDED OR UV-IRRADIATED DOUBLE-STRANDED DNA (BY SIMILARITY).

CC -1- SUBUNIT: CONSISTS OF THREE SUBUNITS; UVRA, UVRB AND UVRC.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. UVRA SUBFAMILY.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

CC

DR EMBL; U34760; AAA84885.1; -.

DR InterPro; IPR003439; ABC\_transportr.

DR InterPro; IPR004602; UVRA.

DR Pfam; PF00005; ABC\_tran; 2.

DR ProDom; PD000006; ABC\_transportr; 1.

DR TIGRFAMs; TIGR00630; uvra; 1.

DR PROSITE; PS00211; ABC\_TRANSPORTER; 2.

DR SOS response; Excision nuclease; DNA repair; ATP-binding; Repeat;

KW DNA-binding; Zinc-finger.

KW NP\_BIND 42 49 ATP (POTENTIAL).

FT NP\_BIND 649 656 ATP (POTENTIAL).

FT ZN\_FING 262 289 C4-TYPE.

FT ZN\_FING 748 774 C4-TYPE.

SO SEQUENCE 950 AA; 105552 MW; 359806240AAD06F1 CRC64;

Query Match 20.0%; Score 6; DB 1; Length 950;

Best Local Similarity 100.0%; Pred. No. 70;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YSLEP 7

Db 267 YSLEP 272

RESULT 29

BCG3\_ACEXY

ID BCG3\_ACEXY STANDARD; PRT; 1325 AA.

AC Q9WK63;

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

```

DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cellulose synthase 1 operon protein C precursor.
GN BCSCI.
OS Acetobacter xylinus.
OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
OC Gluconacetobacter.
OX NCBI_TaxID=28448;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JCM 7664 / IPO 13693;
RX PubMed=10382968;
RA Umeda Y., Hirano A., Ishibashi M., Akiyama H., Onizuka T., Ikeuchi M.,
RA Inoue Y.;
RT "Cloning of cellulose synthase genes from Acetobacter xylinum JCM
RT 7664: implication of a novel set of cellulose synthase genes.";
RL DNA Ref. 6:109-115(1999).
CC -1- FUNCTION: Required for maximal bacterial cellulose synthesis. It
CC may be involved in the formation of a membrane complex for
CC extrusion of the cellulose product (By similarity).
CC -1- PATHWAY: Bacterial cellulose biosynthesis.
CC -1- SUBCELLULAR LOCATION: Outer membrane (Potential).
CC -1- SIMILARITY: BELONGS TO THE ACSC/BCSC FAMILY.
CC -1- SIMILARITY: CONTRAINS 8 TPR REPEATS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AB015802; BAA77587.1; -.
DR InterPro; IPR003921; Cellsynth_C.
DR InterPro; IPR001440; TPR.
DR Pfam; PF00515; TPR; 4.
DR PRINTS; PR01441; CELL5INTHSEC.
KW Cellulose biosynthesis; Repeat; TPR repeat; Membrane; Outer membrane;
KW Signal.
FT SIGNAL 1 30 POTENTIAL.
FT CHAIN 1 1325 CELLULOSE SYNTHASE 1 OPERON PROTEIN C.
FT REPEAT 50 83 TPR 1.
FT REPEAT 85 117 TPR 2.
FT REPEAT 292 325 TPR 3.
FT REPEAT 326 359 TPR 4.
FT REPEAT 406 439 TPR 5.
FT REPEAT 558 591 TPR 6.
FT REPEAT 702 735 TPR 7.
FT REPEAT 737 769 TPR 8.
FT SEQUENCE 1325 AA; 142541 MW; DA4EEGCFADDED74D CRC64;

Query Match 20.0%; Score 6; DB 1; Length 1325;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 18 ISGQIN 23
Db 900 ISGQIN 905

RESULT 30
ID V194 FOMPV STANDARD; PRT; 74 AA.
AC Q95539;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein FV194.
GN FV194.
OS Fowlpox virus (FPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Avipoxvirus.
OX NCBI_TaxID=10261;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20193820; PubMed=10729156;
RA Alfonso C.L., Tulman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RT "The genome of fowlpox virus.";
RL J. Virol. 74:3815-3831(2000).
CC -1- SIMILARITY: BELONGS TO THE POXVIRUSES A30 FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AF198100; AAF4538.1; -.
SQ SEQUENCE 74 AA; 8580 MW; E41581447F94752C CRC64;

Query Match 16.7%; Score 5; DB 1; Length 74;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 10 EFSAT 14
Db 29 EFSAT 33

RESULT 31
ID YKK2 CAEEL STANDARD; PRT; 85 AA.
AC P34279;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein C02F5.2 in chromosome III.
GN C02F5.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainecough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Frazer A.,
RA Johnston L., Jones M., Kershaw J., Kirschen J., Laister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Shownkeen R.,
RA Sims M., Smailson N., Smith A., Smith M., Sonhammer B., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterston R., Watson A., Wellstock L., Wilkinson-Sproat J.,
RA Wohldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; U14745; AAA27917.1; -.
DR PIR; S44604; S44604.
DR WormPep; C02F5.2; CE00038.
KW Hypothetical protein.
SQ SEQUENCE 85 AA; 9425 MW; 10D565D0551A2CCC CRC64;

```

Query Match 16.7%; Score 5; DB 1; Length 85;  
Best Local Similarity 100.0%; Pred. No. 87;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 FSATE 15  
Db 60 FSATE 64

RESULT 32  
YCK5\_CHIRE STANDARD; PRT; 101 AA.  
ID\_YCK5\_CHIRE  
AC P37825;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical 11.4 kDa protein in TRNR-CHLB intergenic region (ORF101).  
OS Chlamydomonas reinhardtii.  
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
OC Chlamydomonadaceae; Chlamydomonas.  
ON NCBI\_TaxID=3055;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=137C / CC-125;  
RX MEDLINE=9403309; PubMed=8219066;  
RA Liu X.-Q., Xu H., Huang C.;  
RT "Chloroplast chl b gene is required for light-independent chlorophyll  
accumulation in Chlamydomonas reinhardtii."  
RL Plant Mol. Biol. 23:297-308(1993).  
KW Chloroplast; Hypothetical protein.  
SQ SEQUENCE 101 AA; 11437 MW; D6B871B65FB3065C CRC64;

Query Match 16.7%; Score 5; DB 1; Length 101;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 ISGOI 22  
Db 26 ISGOI 30

RESULT 33  
Y265\_UREPA STANDARD; PRT; 109 AA.  
ID\_Y265\_UREPA  
AC QPPO6;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein U265.  
GN U265  
OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).  
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Ureaplasma.  
ON NCBI\_TaxID=134821;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Serovar 3;  
RX MEDLINE=20500219; PubMed=11048724;  
RA Glass J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,  
RT "The complete sequence of the mucosal pathogen Ureaplasma  
urealyticum."  
RL Nature 407:757-762(2000).

CC -1- SIMILARITY: BELONGS TO THE UPPO154 FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
CC DR EMBL; AEO02123; AAF30674.1; -.  
CC DR InterPro; IPR005359; UPF0154.  
CC DR Pfam; PF03672; UPF0154; 1.  
CC KW Hypothetical protein; Transmembrane; Complete proteome.  
CC FT TRANSMEM 42 62 POTENTIAL.  
CC SQ SEQUENCE 109 AA; 12223 MW; 00EE24A524356923 CRC64;

Query Match 16.7%; Score 5; DB 1; Length 109;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 QINEL 25  
Db 97 QINEL 101

RESULT 34  
ANFC\_SCYCA STANDARD; PRT; 115 AA.  
ID\_ANFC\_SCYCA  
AC P23259;  
DT 01-NOV-1991 (Rel. 20, Created)  
DT 01-NOV-1991 (Rel. 20, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE C-type natriuretic peptide (CNP-115).  
OS Scyllorhinus canicula (Spotted dogfish) (Spotted catshark).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;  
OC Elasmobranchii; Galeomorphii; Galeoidae; Carcharhiniformes;  
OC Scyllorhinidae; Scyllorhinus.  
ON NCBI\_TaxID=7830;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Heart atrium, and Heart ventricle;  
RX MEDLINE=91243822; PubMed=1828036;  
RA Suzuki R., Takahashi A., Hazon N., Takei Y.;  
RT "Isolation of high-molecular-weight C-type natriuretic peptide from  
the heart of a cartilaginous fish (European dogfish, Scyllorhinus  
canicula)."  
RL FEBS Lett. 282:321-325(1991).

CC -1- FUNCTION: VASORELAXANT ACTIVITY. HAS A GMP-STIMULATING ACTIVITY  
(BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- TISSUE SPECIFICITY: CNP-115 IS DIFFERENTIALLY PROCESSED TO  
CC PRODUCE CNP-38 AND CNP-39 IN THE HEART AND CNP-22 IN THE BRAIN.  
CC -1- SIMILARITY: BELONGS TO THE NATRIURETIC PEPTIDES FAMILY.  
CC PIR; S15822; S15822.  
CC InterPro; IPR002406; C\_natriurtpep.  
CC InterPro; IPR000663; Natr\_peptide.  
CC Pfam; PF00212; ANP; 1.  
CC PRINTS; PR00710; NATPEPTIDES.  
CC ProDom; PD005617; C\_natriurtpep; 1.  
CC SMART; SM00183; NAT\_pep; 1.  
CC PROSITE; PS00263; NATRIURETIC\_PEPTIDE; 1.  
CC KW Vasoactive.  
CC FT PEPTIDE 77 115 CNP-39.  
CC FT PEPTIDE 78 115 CNP-38.  
CC FT PEPTIDE 94 115 CNP-22.  
CC FT DISURFD 99 115 BY SIMILARITY.  
CC SQ SEQUENCE 115 AA; 12885 MW; 49AE7200BE4C7F8A CRC64;

Query Match 16.7%; Score 5; DB 1; Length 115;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SLPEL 7  
Db 40 SLPEL 44

RESULT 35  
SMD2\_SCHPO STANDARD; PRT; 115 AA.  
ID\_SMD2\_SCHPO  
AC O14036;

```

DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable small nuclear ribonucleoprotein Sm D2 (snRNP core protein D2)
DE (Sm-D2).
GN SMD2 OR SPAC2C4.03C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RX MEDLINE=21848401; PubMed=11859360;
RC STRAIN=972;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgourou J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooke K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holtroyd S., Hornby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymoprez B.,
RA Welfens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moesti D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leinbach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Weiler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Leleaux V., Motier S.,
RA Galbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst A.M.,
RA Lucas M., Rochet M., Gallierdin C., Tallada V.A., Garzon A., Rhode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerruti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrett B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -!- FUNCTION: REQUIRED FOR PRE-MRNA SPLICING. REQUIRED FOR SNRNP
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: BELONGS TO THE SNRNP CORE PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z99259; CAB16363.1; -.
CC DR InterPro; IPR001163; snRNP_Sm.
CC DR Pfam; PF01423; Sm; 1.
CC KW Nuclear protein; Ribonucleoprotein; mRNA splicing; mRNA processing.
CC SEQUENCE 115 AA; 13095 MW; EF39B51D94356C33 CRC64;
QY 9 YEFSA 13
DB 23 YEFSA 27

```

```

DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE REV protein (Anti-repression transactivator protein) (ART/TRS).
DE REV.
GN Chimpanzee immunodeficiency virus (SIVcpz) (CIV).
OS Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11723;
RX MEDLINE=90259077; PubMed=2188136;
RC STRAIN=VF5;
RA Hueb T., Cheynier R., Meyerhans A., Roelants G., Wain-Hobson S.;
RT "Genetic Organization of a chimpanzee lentivirus related to HIV-1.";
RL Nature 345:356-359(1990).
CC -!- FUNCTION: REV APPEARS TO ACT POST-TRANSCRIPTIONALLY TO RELIEVE
CC NEGATIVE REPRESSION OF GAG AND ENV PRODUCTION.
CC -!- SUBCELLULAR LOCATION: Nuclear; accumulates in the nucleol1.
CC -!- PTM: PHOSPHOPROTEIN WHOSE STATE OF PHOSPHORYLATION IS MEDIATED
CC BY A SPECIFIC SERINE KINASE ACTIVITY PRESENT IN THE NUCLEUS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X52154; CA36405.1; -.
CC DR PIR; S09988; VKLJST.
CC DR HIV; X52154; REVSCP2.
CC DR InterPro; IPR000625; REV_protein.
CC DR Pfam; PF00424; REV; 1.
CC KW Transcription regulation; AIDS; Phosphorylation; Nuclear protein.
CC SEQUENCE 124 AA; 13701 MW; F5877D1BDF65A7B2 CRC64;
QY 4 LPELD 8
DB 76 LPELD 80

```

```

RESULT 37
ID Y364.AQUAE STANDARD; PRT; 125 AA.
AC 06689;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein AQ_364.
GN AQ_364.
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificae (class); Aquificales; Aquificaceae;
OC Aquifex.
OX NCBI_TaxID=63363;
RX MEDLINE=98196666; PubMed=9537320;
RC STRAIN=VF5;
RA Decker G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aujay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus.";
RL Nature 392:353-358(1998).
CC -!- SIMILARITY: BELONGS TO THE UPF0076 (UKL14) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AE000686; AAC06655.1; -  
DR HSSP; P37552.10D9.  
DR InterPro; IPR000543; YjGF-like.  
DR Pfam; PF01042; UPF0076.1.  
DR TIGRFAMs; TIGR00004; YER057C YjGF\_UK14; 1.  
DR PROSITE; PS01094; UPF0076.1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 125 AA; 14004 MW; 945D559A5C71E819 CRC64;

Query Match 16.7%; Score 5; DB 1; Length 125;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

18 ISGQ1 22  
|||||  
28 ISGQ1 32

RESULT 38  
Y026\_NPVOP  
ID Y026\_NPVOP STANDARD; PRT; 127 AA.  
AC O10297;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE Hypothetical 14.4 kDa protein (ORF42).  
OS Orygia pseudotsugata multicausid polyhedrosis virus (OpnNPV).  
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;  
OC Nucleopolyhedrovirus.  
OX NCBI\_TaxID=164623;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97271300; PubMed=9126251;  
RA Atkins C.H., Russell R.R., Funk C.J., Evans J., Harwood S.,  
RA Rohrmann G.F.;  
RT "The sequence of the Orygia pseudotsugata multicausid nuclear  
RT polyhedrosis virus genome."  
RL Virology 229:381-399(1997).  
CC -1- SIMILARITY: TO CORRESPONDING ORF IN ACNPV.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; U75930; AAC59041.1; -  
KW Hypothetical protein.  
SQ SEQUENCE 127 AA; 14389 MW; 0771366D98517399 CRC64;

Query Match 16.7%; Score 5; DB 1; Length 127;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LDYEF 11  
|||||  
Db 79 LDYEF 83

RESULT 39  
VAL2\_ABMVW  
ID VAL2\_ABMVW STANDARD; PRT; 129 AA.  
AC P21944;  
DT 01-AUG-1991 (Rel. 19, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 01-FEB-1996 (Rel. 33, Last annotation update)

DE AL2 protein.  
GN AC2.  
OS Abutilon mosaic virus (isolate West India).  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=10816;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=9102094; PubMed=2219703;  
RA Frischmuth T., Zilmat G., Jeske H.;  
RT "The nucleotide sequence of abutilon mosaic virus reveals prokaryotic  
RT as well as eukaryotic features."  
RL Virology 178:461-468(1990).  
RN [2]  
RP REVISIONS.  
RA Jeske H.;  
RL Submitted (NOV-1990) to GEMINIVIRUSES AL2 PROTEIN FAMILY.  
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL2 PROTEIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; X15983; CAA34112.1; -  
DR PIR; D36214; OQCW4.  
DR InterPro; IPR000942; Gemin1 AL2.  
DR Pfam; PF01440; Gemin1 AL2; 1.  
DR PRINTS; PR00230; GEMCONVAL2.  
DR ProDom; PD001117; Gemin1 AL2; 1.  
SQ SEQUENCE 129 AA; 14628 MW; 89C4E4DA563B9F CRC64;

Query Match 16.7%; Score 5; DB 1; Length 129;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SLPEL 7  
|||||  
Db 109 SLPEL 113

RESULT 40  
VAL2\_TMOV  
ID VAL2\_TMOV STANDARD; PRT; 129 AA.  
AC Q0658;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 01-JUN-1994 (Rel. 29, Last annotation update)  
DE AL2 protein.  
GN AL2.  
OS Tomato mottle virus (isolate Florida) (TMOV).  
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.  
OX NCBI\_TaxID=36449;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93107858; PubMed=1469361;  
RA Abouzid A.M., Polston J.E., Hiebert E.;  
RT "The nucleotide sequence of tomato mottle virus, a new geminivirus  
RT isolated from tomatoes in Florida."  
RL J. Gen. Virol. 73:3225-3229(1992).  
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL2 PROTEIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; L14460; AAC32417.1; -

DR PIR; J01871; J01871.  
DR InterPro; IPR000942; Gemini AL2.  
DR Pfam; PF01440; Gemini AL2; 1.  
DR PRINTS; PR00230; GEMCOATPA12  
DR ProDom; PD001117; Gemini\_AL2; 1.  
SQ SEQUENCE 129 AA; 14503 MW; AS1FA123739C9D9 CRC64;

Query Match 16.7%; Score 5; DB 1; Length 129;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 SLPEL 7  
|||  
|||  
Db 109 SLPEL 113

Search completed: April 9, 2003, 14:14:55  
Job time : 30.9908 secs

GenCore version 5.1.4 p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 9, 2003, 14:09:35 ; Search time 14.8624 Seconds  
(without alignments)  
415.910 Million cell updates/sec

Title: US-09-987-190-2  
Perfect score: 30  
Sequence: 1 KYSLELDYEFSAATEPYISQINEIXYTX 30

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 671580 seqs, 206047115 residues  
Word size: 0

Total number of hits satisfying chosen parameters: 671580  
Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database:

SPTREMBL\_21.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phage.\*  
10: sp\_plant.\*  
11: sp\_proteic.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_virus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26.7	8	207	2	Q9F9R1
2	26.7	8	210	2	Q9AM00
3	26.7	8	227	3	P79022
4	23.3	7	226	10	Q93X78
5	23.3	7	23.3	10	Q9CSH7
6	23.3	7	462	10	Q942J1
7	23.3	7	495	10	Q91V34
8	23.3	7	23.3	10	Q42975
9	23.3	7	507	10	Q24524
10	23.3	7	23.3	10	Q91V33
11	23.3	7	514	10	Q8VWL8
12	23.3	7	531	10	Q41172
13	23.3	7	541	10	Q40283
14	23.3	7	620	5	Q8SVR3
15	23.3	7	807	5	Q9U275
16	23.3	7	936	2	Q9AMT5

17	7	23.3	1011	8	Q94QB2	Q94qb2 synechococ
18	6	20.0	57	16	Q97AN6	Q97an6 thermoplas
19	6	20.0	91	16	Q820Y8	Q820y8 salmoneila
20	6	20.0	100	15	Q97776	Q97776 human immu
21	6	20.0	150	2	Q9K590	Q9K590 halomonas e
22	6	20.0	153	11	Q9CJ39	Q9Cj39 mus musculi
23	6	20.0	156	5	Q9BPW2	Q9bpw2 metagonimus
24	6	20.0	163	16	Q992V3	Q992v3 streptococ
25	6	20.0	184	1	Q06112	Q06112 methanococ
26	6	20.0	189	2	Q59673	Q59673 propionibac
27	6	20.0	193	15	Q90LU8	Q90lu8 human immu
28	6	20.0	200	2	Q9APR3	Q9ap3 corynebacte
29	6	20.0	200	15	Q9YV92	Q9yv92 human immu
30	6	20.0	200	15	Q90009	Q90009 human immu
31	6	20.0	202	15	Q70722	Q70722 human immu
32	6	20.0	202	15	Q70723	Q70723 human immu
33	6	20.0	202	15	Q71823	Q71823 human immu
34	6	20.0	202	15	Q71824	Q71824 human immu
35	6	20.0	202	15	Q71827	Q71827 human immu
36	6	20.0	203	15	Q90010	Q90010 human immu
37	6	20.0	203	15	Q9YV91	Q9yv91 human immu
38	6	20.0	203	15	Q9YV90	Q9yv90 human immu
39	6	20.0	203	15	Q70724	Q70724 human immu
40	6	20.0	203	15	Q70725	Q70725 human immu
41	6	20.0	203	15	Q70726	Q70726 human immu
42	6	20.0	203	15	Q71825	Q71825 human immu
43	6	20.0	203	15	Q71826	Q71826 human immu
44	6	20.0	205	11	Q91V37	Q91v37 mus musculi
45	6	20.0	206	15	Q70842	Q70842 human immu
46	6	20.0	206	15	Q12256	Q12256 human immu
47	6	20.0	206	15	Q12260	Q12260 human immu
48	6	20.0	206	15	Q12262	Q12262 human immu
49	6	20.0	207	15	Q70835	Q70835 human immu
50	6	20.0	209	12	Q91E19	Q91e19 rice black
51	6	20.0	209	12	Q913E1	Q91e1 rice black
52	6	20.0	209	12	Q910K4	Q910k4 rice black
53	6	20.0	209	12	Q913E3	Q91e3 rice black
54	6	20.0	221	4	Q96T08	Q96t08 homo sapien
55	6	20.0	223	5	Q8SGV7	Q8sgv7 encephalito
56	6	20.0	234	16	Q91656	Q91656 pseudomonas
57	6	20.0	264	15	Q9WPK4	Q9wpk4 human immu
58	6	20.0	264	15	Q9WB86	Q9wb86 human immu
59	6	20.0	264	15	Q9WBQ0	Q9wbq0 human immu
60	6	20.0	266	15	Q9J4Q5	Q9j4q5 human immu
61	6	20.0	283	2	Q31291	Q31291 buchiera ap
62	6	20.0	293	15	Q9J4P0	Q9j4p0 human immu
63	6	20.0	313	16	Q9CTU2	Q9ctu2 pasteurella
64	6	20.0	318	5	Q95YH4	Q95yh4 pseudotrich
65	6	20.0	318	5	Q95YH3	Q95yh3 pseudotrich
66	6	20.0	318	5	Q95NK6	Q95nk6 pseudotrich
67	6	20.0	326	5	Q95P29	Q95p29 holomastigo
68	6	20.0	328	5	Q95P32	Q95p32 pseudotrich
69	6	20.0	338	16	Q9CFF4	Q9cff4 lactococcus
70	6	20.0	340	2	Q68186	Q68186 lactococcus
71	6	20.0	342	16	Q83329	Q83329 treponema p
72	6	20.0	345	5	Q22846	Q22846 caenorhabdi
73	6	20.0	352	17	Q9HP70	Q9hp70 halobacteri
74	6	20.0	354	10	Q9C5W2	Q9c5w2 arbidopsis
75	6	20.0	355	16	Q9HU37	Q9hu37 pseudomonas
76	6	20.0	361	17	Q82XN6	Q82xn6 pyrobaculum
77	6	20.0	366	10	Q9SXX9	Q9sxx9 arbidopsis
78	6	20.0	384	10	Q9SLA0	Q9sla0 arbidopsis
79	6	20.0	385	16	Q9CC15	Q9cc15 arbidopsis
80	6	20.0	385	16	Q05881	Q05881 mycobacteri
81	6	20.0	398	15	Q11594	Q11594 human immu
82	6	20.0	405	2	Q66165	Q66165 agrobacteri
83	6	20.0	405	2	Q52225	Q52225 agrobacteri
84	6	20.0	411	16	Q8U691	Q8u691 agrobacteri
85	6	20.0	412	16	Q97W76	Q97w76 clostridium
86	6	20.0	429	3	Q9UCY9	Q9ucy9 schistosom
87	6	20.0	443	2	Q9ZHL4	Q9zhl4 haemophilus
88	6	20.0	446	16	Q8UBH0	Q8ubh0 agrobacteri
89	6	20.0	448	12	Q91IL0	Q91il0 white spot

90	6	20.0	448	12	Q8QTG4	Q8qtg4 white spot	163	5	16.7	67	16	Q8XPf4	Q8xpF4 clostridium
91	6	20.0	454	6	Q9BDE1	Q9bde1 micromycet	164	5	16.7	68	6	Q9N139	Q9n139 bos taurus
92	6	20.0	454	6	Q9BGC9	Q9bgc9 micromycet	165	5	16.7	68	15	Q97748	Q97748 human immun
93	6	20.0	455	15	Q92MW2	Q92mw2 rhizobium m	166	5	16.7	68	15	Q40491	Q40491 human immun
94	6	20.0	473	15	Q9WRJ3	Q9wrj3 human immun	167	5	16.7	69	2	Q4536	Q4536 azotobacter
95	6	20.0	477	17	Q8ZSN7	Q8zsn7 pyrobaculum	168	5	16.7	69	3	P87091	P87091 cryphonectr
96	6	20.0	482	16	Q97KK6	Q97kk6 clostridium	169	5	16.7	70	5	Q9NMS8	Q9nms8 leishmania
97	6	20.0	487	3	Q9P8R2	Q9p8r2 candida alb	170	5	16.7	70	16	Q9JUH7	Q9juh7 neisseria m
98	6	20.0	488	16	Q9K936	Q9k936 bacillus ha	171	5	16.7	75	6	Q9GMI9	Q9gmi9 macaca fasc
99	6	20.0	496	2	Q69262	Q69262 thermobacil	172	5	16.7	75	12	Q9OB61	Q9ob61 yaba monkey
100	6	20.0	506	10	Q64879	Q64879 arabadopsis	173	5	16.7	75	12	Q9DHJ3	Q9dhj3 yaba-like d
101	6	20.0	507	10	Q9FH03	Q9fho3 arabadopsis	174	5	16.7	76	5	Q2U092	Q2u092 caenorhabdi
102	6	20.0	507	10	Q9LUD0	Q9lud0 arabadopsis	175	5	16.7	77	16	Q8R8H7	Q8r8h7 thermoaer
103	6	20.0	508	16	Q8Z194	Q8z194 yersinia pe	176	5	16.7	79	5	Q9U0Q2	Q9u0q2 plasmodium
104	6	20.0	509	2	Q93F40	Q93f40 shigella fl	177	5	16.7	80	5	Q9U0P4	Q9u0p4 plasmodium
105	6	20.0	511	16	Q8R949	Q8r949 thermoaer	178	5	16.7	80	5	Q9U0P8	Q9u0p8 plasmodium
106	6	20.0	514	10	Q9M1D0	Q9m1d0 arabadopsis	179	5	16.7	80	5	Q9U0P7	Q9u0p7 plasmodium
107	6	20.0	517	10	Q64882	Q64882 arabadopsis	180	5	16.7	80	5	Q9U0P3	Q9u0p3 plasmodium
108	6	20.0	528	10	Q945G7	Q945g7 prunus sero	181	5	16.7	80	5	Q9U0P1	Q9u0p1 plasmodium
109	6	20.0	545	10	Q9M3O9	Q9m3o9 arabadopsis	182	5	16.7	80	5	Q9TVP9	Q9tvp9 plasmodium
110	6	20.0	553	10	Q40984	Q40984 prunus sero	183	5	16.7	80	5	Q9TVN9	Q9tvn9 plasmodium
111	6	20.0	561	10	Q9FPW6	Q9fpw6 arabadopsis	184	5	16.7	82	9	Q8SC60	Q8sc60 plasmodium
112	6	20.0	561	16	Q8XAW0	Q8xaw0 escherichia	185	5	16.7	84	5	Q26111	Q26111 pratylenchu
113	6	20.0	569	8	Q9G401	Q9g401 lithobius f	186	5	16.7	85	16	Q8Y2G9	Q8y2g9 anabaena sp
114	6	20.0	588	16	Q92XR7	Q92xc7 rhizobium m	187	5	16.7	86	2	-Q9ZEM1	Q9zem1 lactobacill1
115	6	20.0	593	2	Q9EY32	Q9ey32 xanthomonas	188	5	16.7	88	9	Q8SC25	Q8sc25 stx2 conver
116	6	20.0	603	8	Q9MR61	Q9mr61 ciconia boy	189	5	16.7	88	16	Q8YWF1	Q8ywf1 anabaena sp
117	6	20.0	617	2	Q68077	Q68077 rhodobacter	190	5	16.7	91	6	Q28857	Q28857 pan troglod
118	6	20.0	620	5	Q9VFS3	Q9vfs3 dirosophila	191	5	16.7	97	6	Q9T5B6	Q9t5b6 crichosurus
119	6	20.0	640	16	Q9C1U9	Q9c1u9 lactococcus	192	5	16.7	99	10	Q9FNC9	Q9fnc9 arabadopsis
120	6	20.0	647	16	Q9WZE0	Q9wze0 thermotoga	193	5	16.7	101	17	Q9YB57	Q9yb57 aeropyrum p
121	6	20.0	650	5	Q9V3X1	Q9v3x1 dirosophila	194	5	16.7	102	15	Q9QWZ1	Q9qwn1 human immun
122	6	20.0	661	16	Q31849	Q31849 bacillus su	195	5	16.7	102	16	Q8RE89	Q8reb9 fusobacteri
123	6	20.0	729	5	Q9VYU1	Q9vyu1 dirosophila	196	5	16.7	103	15	Q78883	Q78883 human immun
124	6	20.0	740	2	Q9P6Z9	Q9p6z9 streptococc	197	5	16.7	103	16	Q8UKJ8	Q8ukj8 agrobacteri
125	6	20.0	779	4	Q96CV5	Q96cv5 homo sapien	198	5	16.7	104	12	Q8VAM3	Q8vam3 white spot
126	6	20.0	785	16	Q9CNO5	Q9cno5 pasteurella	199	5	16.7	104	15	Q69816	Q69816 human immun
127	6	20.0	815	10	Q9SXK0	Q9sxk0 arabadopsis	200	5	16.7	104	15	Q69817	Q69817 human immun
128	6	20.0	855	15	Q902H5	Q902h5 human immun	201	5	16.7	106	5	Q9U6P2	Q9u6p2 plasmodium
129	6	20.0	872	4	Q96165	Q96165 homo sapien	202	5	16.7	106	10	Q9M5V6	Q9m5v6 glycine max
130	6	20.0	879	11	Q9WTP0	Q9wtp0 rattus norv	203	5	16.7	107	15	Q78887	Q78887 human immun
131	6	20.0	899	10	P93698	P93698 vigna ungu	204	5	16.7	107	17	Q9YAA9	Q9yaa9 aeropyrum p
132	6	20.0	948	16	Q9JUS4	Q9juz4 neisseria m	205	5	16.7	108	11	Q9QWD7	Q9qwd7 rattus sp.
133	6	20.0	949	16	Q9JZP1	Q9jzpl neisseria m	206	5	16.7	109	4	Q9BY65	Q9by65 homo sapien
134	6	20.0	954	16	Q8Y2B6	Q8y2b6 ralsconia s	207	5	16.7	109	4	Q9B701	Q9bt01 homo sapien
135	6	20.0	965	16	Q9PAR9	Q9par9 xylella fas	208	5	16.7	110	3	Q9C1J6	Q9c1j6 neurospora
136	6	20.0	1023	5	Q9XYD4	Q9xyd4 dictyosteli	209	5	16.7	110	9	Q9AF61	Q9af61 streptococc
137	6	20.0	1109	16	Q8XMS8	Q8xms8 clostridium	210	5	16.7	111	4	Q9H383	Q9h383 homo sapien
138	6	20.0	1146	3	Q06685	Q06685 saccharomyc	211	5	16.7	112	16	Q9H318	Q9h318 rhizobium 1
139	6	20.0	1190	17	Q8TVN7	Q8tvn7 methanopyru	212	5	16.7	114	15	Q8Q5P6	Q8q5p6 maedi-virna
140	6	20.0	1277	3	Q9C2D2	Q9c2d2 neurospora	213	5	16.7	115	5	Q25836	Q25836 plasmodium
141	6	20.0	1280	2	Q9P6X9	Q9p6x9 chloroflexu	214	5	16.7	115	5	Q9TVW7	Q9tvw7 plasmodium
142	6	20.0	1325	2	Q9WK63	Q9wk63 acetobacter	215	5	16.7	115	11	Q99LQ9	Q99lq9 mus musculu
143	6	20.0	1379	17	Q9HL55	Q9hls5 thermoplasma	216	5	16.7	116	4	Q9N293	Q9nz93 homo sapien
144	6	20.0	1551	11	Q9WTP1	Q9wtp1 rattus norv	217	5	16.7	116	4	Q8WVL9	Q8wvl9 homo sapien
145	6	20.0	1679	3	Q9P7H6	Q9p7h8 schizosacch	218	5	16.7	116	10	Q9SE71	Q9se71 brassica na
146	6	20.0	1687	11	Q35651	Q35651 mus musculu	219	5	16.7	116	10	Q9SE70	Q9se70 brassica na
147	6	20.0	1706	11	P97780	P97780 mus musculu	220	5	16.7	116	10	Q9SE69	Q9se69 brassica ol
148	6	20.0	1719	11	P97789	P97789 mus musculu	221	5	16.7	116	10	Q9SE68	Q9se68 brassica ol
149	6	20.0	1823	5	Q26638	Q26638 paracentrot	222	5	16.7	117	5	Q25797	Q25797 plasmodium
150	6	20.0	2316	2	Q9PDU9	Q9pdj9 bacteroides	223	5	16.7	119	4	Q9BWB3	Q9bwb3 homo sapien
151	6	20.0	3198	5	Q26639	Q26639 strongyloce	224	5	16.7	119	9	Q9EFH0	Q9efh0 homo sapien
152	6	20.0	25	6	Q9N256	Q9n256 pan troglod	225	5	16.7	119	9	Q9XJY4	Q9xjy4 bacterioph
153	6	16.7	25	6	Q9N255	Q9n255 papio hamad	226	5	16.7	119	16	Q8YKJ1	Q8ykj1 anabaena sp
154	6	16.7	25	6	Q9N254	Q9n254 macaca fasc	227	5	16.7	119	16	Q8XK7	Q8xk7 escherichia
155	6	16.7	25	6	Q9N253	Q9n253 macaca mula	228	5	16.7	120	10	Q49990	Q49990 lycopersico
156	6	16.7	32	8	Q9T2Q1	Q9t2q1 pinus sylve	229	5	16.7	120	11	Q9D6Y4	Q9d6y4 mus musculu
157	6	16.7	47	16	Q8XG94	Q8xg94 salmonella	230	5	16.7	121	17	Q8TIR6	Q8tir6 melanosarc
158	6	16.7	49	2	Q9RAU1	Q9rau1 azotobacter	231	5	16.7	122	2	Q05561	Q05561 mycobacteri
159	6	16.7	54	16	Q8Z0J5	Q8z0j5 anabaena sp	232	5	16.7	122	4	Q9BY22	Q9by22 homo sapien
160	6	16.7	61	2	Q57203	Q57203 shigella fl	233	5	16.7	122	10	Q9LMB3	Q9lmb3 arabadopsis
161	6	16.7	61	12	Q91G90	Q91g90 columbid ci	234	5	16.7	122	16	P74159	P74159 synchocyst
162	5	16.7	61	12	Q91G95	Q91g95 columbid ci	235	5	16.7	124	12	Q9DMC7	Q9dmc7 rat cytomeg



236	5	16.7	124	16	Q9RTY7	Q9TEY7	deinococcus	309	5	16.7	149	16	Q8Y7B9	Q8Y7b9	listeria mo
237	5	16.7	124	16	Q8XX03	Q8xx03	ralsctona s	310	5	16.7	150	16	Q9PMT8	Q9pmt8	campylobact
238	5	16.7	125	10	Q9SE72	Q9se72	brassica ca	311	5	16.7	151	16	Q9CKH3	Q9ckh3	pasteurella
239	5	16.7	125	10	Q9SE83	Q9se83	brassica ol	312	5	16.7	152	16	Q55703	Q55703	synecocyst
240	5	16.7	125	16	Q9RSE4	Q9re4	deinococcus	313	5	16.7	154	2	Q51563	Q51563	pseudomonas
241	5	16.7	125	16	Q9Z3S2	Q9z3s2	rhizobium m	314	5	16.7	154	16	Q9KEA5	Q9keas	bacillus ha
242	5	16.7	127	16	Q92MW2	Q92mw2	rhizobium m	315	5	16.7	154	16	Q9ZQ57	Q9zq57	rhizobium m
243	5	16.7	128	6	Q95KM9	Q95km9	ovis aries	316	5	16.7	155	11	Q9D354	Q9d354	mus muscicu
244	5	16.7	128	16	Q07771	Q07771	mycobacteri	317	5	16.7	157	16	Q8R9N1	Q8r9n1	thermoplaea
245	5	16.7	128	5	Q17876	Q17876	caenorhabdi	318	5	16.7	157	17	Q977A5	Q977a5	sulfolobus
246	5	16.7	129	12	Q36623	Q36623	taimo tomat	319	5	16.7	158	15	Q73097	Q73097	human immun
247	5	16.7	129	12	Q71962	Q71962	sida golden	320	5	16.7	158	15	Q73098	Q73098	human immun
248	5	16.7	129	12	Q67561	Q67561	bean dwarf	321	5	16.7	158	15	Q73099	Q73099	human immun
249	5	16.7	129	12	Q96611	Q96611	abutilon mo	322	5	16.7	158	15	Q73100	Q73100	human immun
250	5	16.7	129	12	P89125	P89125	sida golden	323	5	16.7	158	15	Q73101	Q73101	human immun
251	5	16.7	129	12	P89130	P89130	sida golden	324	5	16.7	158	15	Q73102	Q73102	human immun
252	5	16.7	129	12	Q41337	Q41337	tomato leaf	325	5	16.7	158	15	Q73103	Q73103	human immun
253	5	16.7	129	12	Q9YMW6	Q9ymw6	sida golden	326	5	16.7	158	15	Q73105	Q73105	human immun
254	5	16.7	129	12	Q9YMW0	Q9ymw0	sida golden	327	5	16.7	158	15	Q73106	Q73106	human immun
255	5	16.7	129	12	Q91198	Q91198	havana toma	328	5	16.7	158	15	Q73109	Q73109	human immun
256	5	16.7	129	12	Q91204	Q91204	potato yell	329	5	16.7	158	15	Q73110	Q73110	human immun
257	5	16.7	129	12	Q9YL75	Q9yl75	tomato leaf	330	5	16.7	158	15	Q73111	Q73111	human immun
258	5	16.7	132	2	Q33019	Q33019	mycobacteri	331	5	16.7	158	15	Q73112	Q73112	human immun
259	5	16.7	132	12	Q913X7	Q913x7	pepper huas	332	5	16.7	158	15	Q73113	Q73113	human immun
260	5	16.7	134	16	Q98K58	Q98k58	rhizobium l	333	5	16.7	158	15	Q73114	Q73114	human immun
261	5	16.7	134	16	Q8YU00	Q8ym00	anabaena sp	334	5	16.7	158	15	Q73116	Q73116	human immun
262	5	16.7	135	10	Q9FYR8	Q9fyR8	eucommia ul	335	5	16.7	158	15	Q73121	Q73121	human immun
263	5	16.7	135	16	Q9ACD6	Q9acd6	caulobacter	336	5	16.7	159	2	Q07466	Q07466	rhodospseudo
264	5	16.7	135	16	Q8VAA5	Q8vaa5	listeria mo	337	5	16.7	159	15	Q9JBT1	Q9jbt1	human immun
265	5	16.7	136	2	Q9A112	Q9a112	photorhabdu	338	5	16.7	159	15	Q9JEBQ	Q9jebq	human immun
266	5	16.7	137	6	Q97509	Q97509	capria hircu	339	5	16.7	159	15	Q9JEBQ3	Q9jebq3	human immun
267	5	16.7	137	12	Q55581	Q55581	leucania se	340	5	16.7	159	16	Q53709	Q53709	mycobacteri
268	5	16.7	137	13	Q9YHH6	Q9yhh6	ginglymysto	341	5	16.7	160	12	Q8QTF4	Q8qtf4	white spot
269	5	16.7	137	16	Q06257	Q06257	strepptomyc	342	5	16.7	160	15	Q73104	Q73104	human immun
270	5	16.7	138	4	Q9NV55	Q9nv55	homo sapien	343	5	16.7	160	15	Q73107	Q73107	human immun
271	5	16.7	139	13	Q9YHN8	Q9yhn8	ginglymysto	344	5	16.7	160	15	Q73108	Q73108	human immun
272	5	16.7	141	9	Q8SC44	Q8sc44	etx2 conyer	345	5	16.7	160	15	Q73115	Q73115	human immun
273	5	16.7	141	12	Q9QRZ9	Q9qrz9	human rotav	346	5	16.7	160	15	Q73118	Q73118	human immun
274	5	16.7	141	12	Q9QRZ8	Q9qrz8	human rotav	347	5	16.7	160	15	Q73122	Q73122	human immun
275	5	16.7	141	12	Q9QRZ6	Q9qrz6	human rotav	348	5	16.7	161	11	Q61430	Q61430	human immun
276	5	16.7	141	12	Q9QRZ5	Q9qrz5	human rotav	349	5	16.7	162	12	Q8YBB3	Q8ybb3	white spot
277	5	16.7	141	12	Q9QRZ4	Q9qrz4	human rotav	350	5	16.7	162	16	Q9CV98	Q9cv98	lactococcus
278	5	16.7	141	12	Q9QRZ3	Q9qrz3	human rotav	351	5	16.7	162	16	Q9ZSE8	Q9zse8	rhizobium m
279	5	16.7	141	12	Q9QRZ2	Q9qrz2	human rotav	352	5	16.7	164	2	Q9JPD3	Q9jpd3	rhodocyclus
280	5	16.7	141	12	Q9QRZ1	Q9qrz1	human rotav	353	5	16.7	164	5	Q9V7H5	Q9v7h5	white spot
281	5	16.7	141	12	Q9QRZ0	Q9qrz0	human rotav	354	5	16.7	164	6	P79252	P79252	oryctolagus
282	5	16.7	141	12	Q9QRX9	Q9qrx9	human rotav	355	5	16.7	165	2	Q45403	Q45403	bacillus su
283	5	16.7	141	12	Q9QRX8	Q9qrx8	human rotav	356	5	16.7	165	16	P73153	P73153	synecocyst
284	5	16.7	141	12	Q9QRX7	Q9qrx7	human rotav	357	5	16.7	165	16	Q8Y158	Q8y158	ralsctonia s
285	5	16.7	141	12	Q9QRX6	Q9qrx6	human rotav	358	5	16.7	166	9	Q03915	Q03915	dactylopha
286	5	16.7	141	12	Q9QRX5	Q9qrx5	human rotav	359	5	16.7	166	10	Q91867	Q91867	arabidopsis
287	5	16.7	141	12	Q9QRX4	Q9qrx4	human rotav	360	5	16.7	168	10	Q9FYI3	Q9fyi3	arabidopsis
288	5	16.7	141	12	Q9QRX3	Q9qrx3	human rotav	361	5	16.7	168	16	Q9RMQ0	Q9rmq0	delinococcus
289	5	16.7	141	12	Q9DKS3	Q9dks3	human rotav	362	5	16.7	168	16	Q9K495	Q9k495	strepptomyc
290	5	16.7	141	16	Q8X5E2	Q8x5e2	escherichia	363	5	16.7	169	5	Q9NFM1	Q9nfm1	schistosoma
291	5	16.7	142	8	Q9TJ75	Q9tj75	gretum grem	364	5	16.7	169	11	Q9CRF9	Q9crf9	mus muscicu
292	5	16.7	142	12	Q9OS00	Q9os00	human rotav	365	5	16.7	169	11	Q89060	Q89060	mus muscicu
293	5	16.7	142	12	Q9QRX2	Q9qrx2	human rotav	366	5	16.7	169	15	Q9E2U6	Q9e2u6	human immun
294	5	16.7	142	12	Q9QRX1	Q9qrx1	human rotav	367	5	16.7	169	16	Q9CK49	Q9ck49	pasteurella
295	5	16.7	142	12	Q9QRX0	Q9qrx0	human rotav	368	5	16.7	169	17	Q29109	Q29109	archaeoglob
296	5	16.7	142	12	Q9QRX9	Q9qrx9	human rotav	369	5	16.7	170	4	Q8WYJ4	Q8wyj4	homo sapien
297	5	16.7	143	2	Q9AKY3	Q9aky3	legionella	370	5	16.7	170	5	Q95PB1	Q95pb1	bombyx mori
298	5	16.7	143	10	Q9AMZ9	Q9amz9	gulliardia	371	5	16.7	171	12	Q9QAU8	Q9qau8	human rotav
299	5	16.7	144	4	Q9N082	Q9n082	homo sapien	372	5	16.7	171	12	Q9QAU7	Q9qau7	human rotav
300	5	16.7	145	5	Q9N3V4	Q9n3v4	caenorhabdi	373	5	16.7	171	12	Q9QAU6	Q9qau6	human rotav
301	5	16.7	145	8	Q9WFB3	Q9wfb3	beta vulgar	374	5	16.7	171	12	Q9QAU4	Q9qau4	human rotav
302	5	16.7	145	13	Q9YH09	Q9yh09	ginglymysto	375	5	16.7	171	12	Q9QAU3	Q9qau3	human rotav
303	5	16.7	147	16	Q8VKH8	Q8vkh8	mycobacteri	376	5	16.7	171	12	Q9QAU2	Q9qau2	human rotav
304	5	16.7	147	17	Q58828	Q58828	pyrococcus	377	5	16.7	171	12	Q9QAU1	Q9qau1	human rotav
305	5	16.7	147	17	Q9UZL1	Q9uzl1	pyrococcus	378	5	16.7	171	12	Q9QAU0	Q9qau0	human rotav
306	5	16.7	149	2	Q9ZEV1	Q9zev1	bacillus st	379	5	16.7	171	12	Q9QAT9	Q9qat9	human rotav
307	5	16.7	149	5	Q8S286	Q8s286	drosophila	380	5	16.7	171	12	Q9QAT8	Q9qat8	human rotav
308	5	16.7	149	16	Q92BY8	Q92by8	listeria in	381	5	16.7	171	12	Q9QAT7	Q9qat7	human rotav

382	5	16.7	171	12	Q9QAT6	Q9qat6 human rotav	455	5	16.7	198	15	Q9E4G7	Q9e4g7 human immun
383	5	16.7	172	8	Q959YE	Q959ye kauplichthys	456	5	16.7	198	15	Q9E4G8	Q9e4g8 human immun
384	5	16.7	173	8	Q8VVE3	Q8vve3 thermus the	457	5	16.7	199	6	Q95WQ1	Q95wq1 bos taurus
385	5	16.7	173	4	Q96LZ3	Q96lcz homo sapien	458	5	16.7	199	15	Q9JDM8	Q9jdm8 human immun
386	5	16.7	173	11	Q9D6E2	Q9d6e2 mus musculu	459	5	16.7	199	15	Q9JDM4	Q9jdm4 human immun
387	5	16.7	174	16	Q8YCC3	Q8ycc3 bruceia me	460	5	16.7	199	15	Q9JDM0	Q9jdm0 human immun
388	5	16.7	175	8	Q99605	Q99605 artibeus ja	461	5	16.7	199	15	Q9JDM8	Q9jdm8 human immun
389	5	16.7	175	16	Q9RMZ3	Q9rmz3 deinococcus	462	5	16.7	199	15	Q9JDM7	Q9jdm7 human immun
390	5	16.7	176	17	Q9HK81	Q9hk81 thermoplasma	463	5	16.7	199	15	Q9JDM4	Q9jdm4 human immun
391	5	16.7	176	17	Q975U3	Q975u3 sulfobobus	464	5	16.7	199	15	Q9JDM2	Q9jdm2 human immun
392	5	16.7	178	12	Q9WNT3	Q9wnt3 trichoplusia	465	5	16.7	199	15	Q9JDM1	Q9jdm1 human immun
393	5	16.7	178	15	Q706I0	Q706i0 human immun	466	5	16.7	199	15	Q9JDM6	Q9jdm6 human immun
394	5	16.7	178	15	Q706I1	Q706i1 human immun	467	5	16.7	199	15	Q9JDM5	Q9jdm5 human immun
395	5	16.7	179	11	Q980MS	Q980ms sulfobobus	468	5	16.7	199	15	Q9JDM0	Q9jdm0 human immun
396	5	16.7	179	16	Q92JH6	Q92jhe rickettsia	469	5	16.7	199	15	Q9JDS9	Q9jds9 human immun
397	5	16.7	181	5	Q8SMCI	Q8smci encephalito	470	5	16.7	199	15	Q9JDM1	Q9jdm1 human immun
398	5	16.7	181	11	Q921W1	Q921w1 mus musculu	471	5	16.7	199	15	Q9JDM5	Q9jdm5 human immun
399	5	16.7	183	5	Q16043	Q16043 drosophila	472	5	16.7	199	15	Q9JDM9	Q9jdm9 human immun
400	5	16.7	183	15	Q70609	Q70609 human immun	473	5	16.7	199	15	Q9JDM8	Q9jdm8 human immun
401	5	16.7	183	15	Q70614	Q70614 human immun	474	5	16.7	199	15	Q9JDM6	Q9jdm6 human immun
402	5	16.7	183	15	Q70616	Q70616 human immun	475	5	16.7	199	15	Q9JDM5	Q9jdm5 human immun
403	5	16.7	183	15	Q70617	Q70617 human immun	476	5	16.7	199	15	Q9JDM4	Q9jdm4 human immun
404	5	16.7	183	15	Q70621	Q70621 human immun	477	5	16.7	199	15	Q9JDM3	Q9jdm3 human immun
405	5	16.7	183	15	Q8UQ86	Q8uq86 human immun	478	5	16.7	199	15	Q9JDM2	Q9jdm2 human immun
406	5	16.7	183	15	Q8UQ84	Q8uq84 human immun	479	5	16.7	199	15	Q9JDM1	Q9jdm1 human immun
407	5	16.7	183	15	Q8UQ83	Q8uq83 human immun	480	5	16.7	199	15	Q9JDM0	Q9jdm0 human immun
408	5	16.7	183	15	Q8UQ82	Q8uq82 human immun	481	5	16.7	199	15	Q9JDM9	Q9jdm9 human immun
409	5	16.7	183	15	Q8UQ81	Q8uq81 human immun	482	5	16.7	199	15	Q9JDM4	Q9jdm4 human immun
410	5	16.7	183	15	Q8UQ80	Q8uq80 human immun	483	5	16.7	199	15	Q9JDM3	Q9jdm3 human immun
411	5	16.7	183	15	Q8UQ79	Q8uq79 human immun	484	5	16.7	199	15	Q9JDM2	Q9jdm2 human immun
412	5	16.7	183	15	Q8UQ78	Q8uq78 human immun	485	5	16.7	199	15	Q9JBM7	Q9jbm7 human immun
413	5	16.7	183	15	Q8UQ77	Q8uq77 human immun	486	5	16.7	199	15	Q8UQ94	Q8uq94 human immun
414	5	16.7	183	15	Q8UQ76	Q8uq76 human immun	487	5	16.7	200	10	Q80711	Q80711 arabidopsis
415	5	16.7	183	15	Q8UQ75	Q8uq75 human immun	488	5	16.7	200	15	Q90008	Q90008 human immun
416	5	16.7	183	15	Q8UQ74	Q8uq74 human immun	489	5	16.7	200	15	Q9YV86	Q9yv86 human immun
417	5	16.7	183	15	Q8UQ73	Q8uq73 human immun	490	5	16.7	200	15	Q9JDM6	Q9jdm6 human immun
418	5	16.7	183	15	Q8UQ72	Q8uq72 human immun	491	5	16.7	200	15	Q9JDM5	Q9jdm5 human immun
419	5	16.7	183	15	Q8UQ70	Q8uq70 human immun	492	5	16.7	200	15	Q9JDM3	Q9jdm3 human immun
420	5	16.7	183	15	Q8UQ69	Q8uq69 human immun	493	5	16.7	200	15	Q9JDM2	Q9jdm2 human immun
421	5	16.7	183	15	Q8UQ68	Q8uq68 human immun	494	5	16.7	200	15	Q9JDM0	Q9jdm0 human immun
422	5	16.7	183	15	Q8UQ67	Q8uq67 human immun	495	5	16.7	200	15	Q9JDM9	Q9jdm9 human immun
423	5	16.7	183	15	Q8UQ65	Q8uq65 human immun	496	5	16.7	201	3	Q9P418	Q9p418 coccidioides
424	5	16.7	183	15	Q8UQ64	Q8uq64 human immun	497	5	16.7	201	15	Q90031	Q90031 human immun
425	5	16.7	183	15	Q8UQ63	Q8uq63 human immun	498	5	16.7	201	15	Q9YRW7	Q9yrw7 human immun
426	5	16.7	183	17	Q29155	Q29155 archaeoglob	499	5	16.7	201	15	Q9YRW3	Q9yrw3 human immun
427	5	16.7	185	2	Q44518	Q44518 anabaena va	500	5	16.7	201	15	Q9JDM8	Q9jdm8 human immun
428	5	16.7	185	16	Q8YV89	Q8yv89 anabaena sp	501	5	16.7	201	15	Q9JDM7	Q9jdm7 human immun
429	5	16.7	186	2	Q52994	Q52994 escherichia	502	5	16.7	201	15	Q9JDM6	Q9jdm6 human immun
430	5	16.7	187	16	Q8Y061	Q8y061 ralestonia s	503	5	16.7	201	15	Q9JDM4	Q9jdm4 human immun
431	5	16.7	188	4	Q96J67	Q96j67 homo sapien	504	5	16.7	201	15	Q9JDM2	Q9jdm2 human immun
432	5	16.7	188	12	Q90F44	Q90f44 avian pneum	505	5	16.7	201	15	Q9JDM0	Q9jdm0 human immun
433	5	16.7	189	3	Q94694	Q94694 schizosacch	506	5	16.7	201	15	Q8UQ47	Q8uq47 human immun
434	5	16.7	189	16	Q8Y1A0	Q8y1a0 ralestonia s	507	5	16.7	201	15	Q8UQ45	Q8uq45 human immun
435	5	16.7	192	16	Q8RH38	Q8rh38 fusobacteri	508	5	16.7	201	15	Q8UQ44	Q8uq44 human immun
436	5	16.7	193	2	Q9ZGN1	Q9zgn1 azotobacter	509	5	16.7	201	15	Q8UQ43	Q8uq43 human immun
437	5	16.7	193	5	Q9VKQ8	Q9vkq8 drosophila	510	5	16.7	201	15	Q8UQ42	Q8uq42 human immun
438	5	16.7	194	15	Q9OLV2	Q9olv2 human immun	511	5	16.7	201	15	Q8UQ41	Q8uq41 human immun
439	5	16.7	194	16	Q8UBS8	Q8ubs8 agrobacteri	512	5	16.7	201	15	Q8UQ40	Q8uq40 human immun
440	5	16.7	195	11	Q8RS12	Q8rs12 mus musculu	513	5	16.7	201	15	Q8UQ39	Q8uq39 human immun
441	5	16.7	195	16	Q51524	Q51524 borelia bu	514	5	16.7	201	15	Q8UQ98	Q8uq98 human immun
442	5	16.7	195	16	P73665	P73665 synechocyst	515	5	16.7	201	15	Q8UQ97	Q8uq97 human immun
443	5	16.7	196	15	Q8QDC7	Q8qdc7 human immun	516	5	16.7	201	15	Q8UQ96	Q8uq96 human immun
444	5	16.7	196	16	Q99T18	Q99t18 staphylococ	517	5	16.7	201	15	Q8UQ95	Q8uq95 human immun
445	5	16.7	196	16	Q8YQ49	Q8yq49 anabaena sp	518	5	16.7	201	15	Q8UQ92	Q8uq92 human immun
446	5	16.7	197	10	Q93WS2	Q93ws2 davidia inv	519	5	16.7	201	15	Q8UQ91	Q8uq91 human immun
447	5	16.7	197	15	Q9JDQ1	Q9jdd1 human immun	520	5	16.7	201	15	Q8UQ89	Q8uq89 human immun
448	5	16.7	197	15	Q9JDP7	Q9jdp7 human immun	521	5	16.7	201	15	Q8UQ88	Q8uq88 human immun
449	5	16.7	197	15	Q9JDN7	Q9jdn7 human immun	522	5	16.7	201	15	Q8UQ87	Q8uq87 human immun
450	5	16.7	198	15	Q9JDN1	Q9jdn1 human immun	523	5	16.7	201	15	Q8QD75	Q8qd75 human immun
451	5	16.7	198	15	Q71835	Q71835 human immun	524	5	16.7	202	8	Q47094	Q47094 erigia asia
452	5	16.7	198	15	Q90409	Q90409 human immun	525	5	16.7	202	15	Q89996	Q89996 human immun
453	5	16.7	198	15	Q9E4G2	Q9e4g2 human immun	526	5	16.7	202	15	Q90003	Q90003 human immun
454	5	16.7	198	15	Q9E4G2	Q9e4g2 human immun	527	5	16.7	202	15	Q90024	Q90024 human immun

528	5	16.7	202	15	090030	090030 human	immun	601	5	16.7	204	15	090408	090408 human	immun
529	5	16.7	202	15	090032	090032 human	immun	602	5	16.7	204	15	090412	090412 human	immun
530	5	16.7	202	15	09YRW9	09YRW9 human	immun	603	5	16.7	204	15	090401	090401 human	immun
531	5	16.7	202	15	09JDX0	09JDX0 human	immun	604	5	16.7	204	15	0904T9	0904T9 human	immun
532	5	16.7	202	15	09JDX5	09JDX5 human	immun	605	5	16.7	204	15	09E4H5	09E4H5 human	immun
533	5	16.7	202	15	09JDX4	09JDX4 human	immun	606	5	16.7	205	5	08S8T0	08S8T0	
534	5	16.7	202	15	09JDX3	09JDX3 human	immun	607	5	16.7	205	8	047093	047093	
535	5	16.7	202	15	09JDX1	09JDX1 human	immun	608	5	16.7	205	15	09WL18	09WL18 human	immun
536	5	16.7	202	15	09JDX0	09JDX0 human	immun	609	5	16.7	205	15	09WL17	09WL17 human	immun
537	5	16.7	202	15	09JDX6	09JDX6 human	immun	610	5	16.7	205	15	0904U5	0904U5 human	immun
538	5	16.7	202	15	09JDX5	09JDX5 human	immun	611	5	16.7	205	16	09PB02	09PB02	
539	5	16.7	202	15	09JDX3	09JDX3 human	immun	612	5	16.7	205	16	09S8E4	09S8E4	
540	5	16.7	202	15	09JDX7	09JDX7 human	immun	613	5	16.7	206	2	08RPM8	08RPM8	
541	5	16.7	202	15	09JDX6	09JDX6 human	immun	614	5	16.7	206	3	096UT6	096UT6	
542	5	16.7	202	15	09JDX0	09JDX0 human	immun	615	5	16.7	206	15	0904U3	0904U3 human	immun
543	5	16.7	202	15	09JDX3	09JDX3 human	immun	616	5	16.7	206	15	08ODD7	08ODD7 human	immun
544	5	16.7	202	15	09E4V8	09E4V8 human	immun	617	5	16.7	206	15	08ODA1	08ODA1 human	immun
545	5	16.7	202	15	09E4W0	09E4W0 human	immun	618	5	16.7	206	15	08QD89	08QD89 human	immun
546	5	16.7	202	15	09E4W1	09E4W1 human	immun	619	5	16.7	206	15	08QD87	08QD87 human	immun
547	5	16.7	202	15	09E4W4	09E4W4 human	immun	620	5	16.7	206	17	09HSV9	09HSV9	
548	5	16.7	202	15	09E4W7	09E4W7 human	immun	621	5	16.7	207	8	021512	021512	
549	5	16.7	202	15	09E4W9	09E4W9 human	immun	622	5	16.7	207	15	09JBA7	09JBA7 human	immun
550	5	16.7	202	15	09E4X0	09E4X0 human	immun	623	5	16.7	207	15	09JB45	09JB45 human	immun
551	5	16.7	202	15	09E4X1	09E4X1 human	immun	624	5	16.7	207	15	09B4C5	09B4C5 human	immun
552	5	16.7	202	15	09E4X2	09E4X2 human	immun	625	5	16.7	207	15	09B4D0	09B4D0 human	immun
553	5	16.7	202	15	09E4X3	09E4X3 human	immun	626	5	16.7	207	15	09E4D7	09E4D7 human	immun
554	5	16.7	202	15	09E4X4	09E4X4 human	immun	627	5	16.7	207	15	09E4B9	09E4B9 human	immun
555	5	16.7	202	15	09E4X5	09E4X5 human	immun	628	5	16.7	207	15	09E4F0	09E4F0 human	immun
556	5	16.7	202	15	09E4X7	09E4X7 human	immun	629	5	16.7	207	15	09E4F1	09E4F1 human	immun
557	5	16.7	202	15	09Y112	09Y112 human	immun	630	5	16.7	207	15	09E4F2	09E4F2 human	immun
558	5	16.7	202	15	09Y112	09Y112 human	immun	631	5	16.7	207	15	09E4F3	09E4F3 human	immun
559	5	16.7	203	6	028682	028682 oryctolagus		632	5	16.7	207	15	09E4F6	09E4F6 human	immun
560	5	16.7	203	6	029032	029032 sus scrofa		633	5	16.7	207	15	09E4F7	09E4F7 human	immun
561	5	16.7	203	15	090021	090021 human	immun	634	5	16.7	207	15	09E4G0	09E4G0 human	immun
562	5	16.7	203	15	09YRX3	09YRX3 human	immun	635	5	16.7	207	15	09E4G9	09E4G9 human	immun
563	5	16.7	203	15	09YRX2	09YRX2 human	immun	636	5	16.7	207	15	09E4H0	09E4H0 human	immun
564	5	16.7	203	15	09YRX1	09YRX1 human	immun	637	5	16.7	207	15	09E4H4	09E4H4 human	immun
565	5	16.7	203	15	09YRX0	09YRX0 human	immun	638	5	16.7	207	15	09E4H6	09E4H6 human	immun
566	5	16.7	203	15	09YRW8	09YRW8 human	immun	639	5	16.7	207	15	08QDD5	08QDD5 human	immun
567	5	16.7	203	15	09YRW6	09YRW6 human	immun	640	5	16.7	207	15	08QDD4	08QDD4 human	immun
568	5	16.7	203	15	09YRW5	09YRW5 human	immun	641	5	16.7	207	15	08QDD0	08QDD0 human	immun
569	5	16.7	203	15	09YRW4	09YRW4 human	immun	642	5	16.7	207	15	08QDC4	08QDC4 human	immun
570	5	16.7	203	15	09YRW2	09YRW2 human	immun	643	5	16.7	207	15	08QDB3	08QDB3 human	immun
571	5	16.7	203	15	09YRW1	09YRW1 human	immun	644	5	16.7	207	15	08QDA6	08QDA6 human	immun
572	5	16.7	203	15	09JDR5	09JDR5 human	immun	645	5	16.7	207	15	08QDA0	08QDA0 human	immun
573	5	16.7	203	15	09JDR3	09JDR3 human	immun	646	5	16.7	207	15	08QD99	08QD99 human	immun
574	5	16.7	203	15	09JDN5	09JDN5 human	immun	647	5	16.7	207	15	08QD88	08QD88 human	immun
575	5	16.7	203	15	09Q007	09Q007 human	immun	648	5	16.7	207	15	08QD85	08QD85 human	immun
576	5	16.7	203	15	071834	071834 human	immun	649	5	16.7	207	15	08QD84	08QD84 human	immun
577	5	16.7	203	15	0904U7	0904U7 human	immun	650	5	16.7	207	15	08QD82	08QD82 human	immun
578	5	16.7	203	15	0904U6	0904U6 human	immun	651	5	16.7	207	15	08QD73	08QD73 human	immun
579	5	16.7	203	15	09E4V1	09E4V1 human	immun	652	5	16.7	207	15	08QD71	08QD71 human	immun
580	5	16.7	203	15	09E4V8	09E4V8 human	immun	653	5	16.7	207	15	08QD69	08QD69 human	immun
581	5	16.7	203	15	09YUG1	09YUG1 human	immun	654	5	16.7	207	15	08Q7J9	08Q7J9 human	immun
582	5	16.7	203	15	09YUG1	09YUG1 human	immun	655	5	16.7	207	16	031588	031588	
583	5	16.7	204	8	047095	047095 striga gesn		656	5	16.7	207	17	08TN94	08TN94	
584	5	16.7	204	10	040023	040023 bordeum vul		657	5	16.7	207	17	08TN84	08TN84	
585	5	16.7	204	12	039496	039496 bovine herp		658	5	16.7	208	4	09EPM5	09EPM5	
586	5	16.7	204	12	065545	065545 bovine herp		659	5	16.7	208	6	097739	097739	
587	5	16.7	204	15	089985	089985 human	immun	660	5	16.7	208	8	021506	021506	
588	5	16.7	204	15	09WLJ0	09WLJ0 human	immun	661	5	16.7	208	15	089987	089987 human	immun
589	5	16.7	204	15	090011	090011 human	immun	662	5	16.7	208	15	089995	089995 human	immun
590	5	16.7	204	15	09JDN6	09JDN6 human	immun	663	5	16.7	208	15	090025	090025 human	immun
591	5	16.7	204	15	09JDN9	09JDN9 human	immun	664	5	16.7	208	15	09E4D2	09E4D2 human	immun
592	5	16.7	204	15	09JDN8	09JDN8 human	immun	665	5	16.7	208	15	08QD98	08QD98 human	immun
593	5	16.7	204	15	09JDN7	09JDN7 human	immun	666	5	16.7	208	15	08QD79	08QD79 human	immun
594	5	16.7	204	15	09JDN6	09JDN6 human	immun	667	5	16.7	208	17	028099	028099	
595	5	16.7	204	15	09JBN2	09JBN2 human	immun	668	5	16.7	209	12	08MPK7	08MPK7	
596	5	16.7	204	15	09JBN0	09JBN0 human	immun	669	5	16.7	209	15	055650	055650	
597	5	16.7	204	15	09JBM7	09JBM7 human	immun	670	5	16.7	209	15	08QDD3	08QDD3 human	immun
598	5	16.7	204	15	09JBM5	09JBM5 human	immun	671	5	16.7	209	15	08QD77	08QD77 human	immun
599	5	16.7	204	15	09JBM1	09JBM1 human	immun	672	5	16.7	209	15	08QAW1	08QAW1 human	immun
600	5	16.7	204	15	09JBL9	09JBL9 human	immun	673	5	16.7	209	15	08QAV8	08QAV8 human	immun

674	5	16.7	209	16	Q97KW0	Q97KW0 clostridium	747	5	16.7	222	16	Q54191	Q54191 streptomyces
675	5	16.7	209	17	Q8TQ69	Q8TQ69 methanosaarc	748	5	16.7	223	16	Q9HUS1	Q9HUS1 pseudomonas
676	5	16.7	210	6	Q62813	Q62813 papio hamad	749	5	16.7	223	17	Q8TM30	Q8TM30 methanosaarc
677	5	16.7	210	6	Q28273	Q28273 canis famli	750	5	16.7	224	2	Q46369	Q46369 chlorobium
678	5	16.7	210	15	Q9JDR7	Q9JDR7 human immu	751	5	16.7	224	2	Q8RJ77	Q8RJ77 nitrosococ
679	5	16.7	210	15	Q8QAU4	Q8QAU4 human immu	752	5	16.7	224	4	Q00617	Q00617 homo sapien
680	5	16.7	210	15	Q8QAU1	Q8QAU1 human immu	753	5	16.7	224	6	Q28848	Q28848 smintropsis
681	5	16.7	210	16	Q9KNI2	Q9KNI2 vibrio chol	754	5	16.7	225	4	Q9BRQ7	Q9BRQ7 homo sapien
682	5	16.7	210	17	Q9JUG8	Q9JUG8 sulfolobus	755	5	16.7	225	5	Q17715	Q17715 caenorhabdi
683	5	16.7	211	11	Q9DIW6	Q9DIW6 mus musculu	756	5	16.7	225	5	Q8STN5	Q8STN5 encephalito
684	5	16.7	211	12	Q39300	Q39300 equine hecp	757	5	16.7	225	10	Q8RYW7	Q8RYW7 oryza sativ
685	5	16.7	211	15	Q8Q7K1	Q8Q7K1 human immu	758	5	16.7	225	12	Q9PYW4	Q9PYW4 xestia c-ni
686	5	16.7	211	16	Q9HYZ4	Q9HYZ4 pseudomonas	759	5	16.7	225	15	Q8QAT3	Q8QAT3 human immu
687	5	16.7	211	16	Q96Y84	Q96Y84 sulfilobus	760	5	16.7	225	17	Q9UZP3	Q9UZP3 pyrococcus
688	5	16.7	212	6	Q28512	Q28512 macaca mula	761	5	16.7	226	5	Q9NHB3	Q9NHB3 mytilus edu
689	5	16.7	212	6	Q28567	Q28567 ovis aries	762	5	16.7	226	16	Q8YXV7	Q8YXV7 anabaena sp
690	5	16.7	212	10	Q9ZW80	Q9ZW80 arabidopsis	763	5	16.7	227	2	Q9S503	Q9S503 myxococcus
691	5	16.7	213	15	Q77609	Q77609 human immu	764	5	16.7	227	11	P70103	P70103 cynomys lud
692	5	16.7	213	15	Q77616	Q77616 human immu	765	5	16.7	228	15	Q89550	Q89550 human immu
693	5	16.7	213	15	Q77629	Q77629 human immu	766	5	16.7	228	16	Q97DC0	Q97DC0 clostridium
694	5	16.7	213	15	Q77630	Q77630 human immu	767	5	16.7	229	16	Q8Y2R4	Q8Y2R4 ralstonia s
695	5	16.7	213	15	Q77651	Q77651 human immu	768	5	16.7	229	16	Q93MA2	Q93MA2 clostridia
696	5	16.7	213	15	Q77653	Q77653 human immu	769	5	16.7	230	2	Q9K570	Q9K570 lactococcus
697	5	16.7	213	15	Q77675	Q77675 human immu	770	5	16.7	230	15	Q9E052	Q9E052 human immu
698	5	16.7	213	15	Q8QDC6	Q8QDC6 human immu	771	5	16.7	230	16	P74631	P74631 synchocyst
699	5	16.7	214	15	Q77620	Q77620 human immu	772	5	16.7	230	16	Q9CF87	Q9CF87 lactococcus
700	5	16.7	214	15	Q77685	Q77685 human immu	773	5	16.7	231	11	Q9D4Q8	Q9D4Q8 mus musculu
701	5	16.7	214	15	Q79367	Q79367 human immu	774	5	16.7	231	15	Q56309	Q56309 wallaye. epi
702	5	16.7	214	15	Q79372	Q79372 human immu	775	5	16.7	231	15	Q9RV00	Q9RV00 delnecoccus
703	5	16.7	214	15	Q79375	Q79375 human immu	776	5	16.7	231	16	Q8X8L2	Q8X8L2 escherichia
704	5	16.7	214	15	Q79376	Q79376 human immu	777	5	16.7	231	16	Q8UBZ0	Q8UBZ0 agrobacteri
705	5	16.7	214	15	Q79380	Q79380 human immu	778	5	16.7	231	16	Q92PB8	Q92PB8 rhizobium m
706	5	16.7	214	15	Q79381	Q79381 human immu	779	5	16.7	232	11	Q99P51	Q99P51 mus musculu
707	5	16.7	214	15	Q79382	Q79382 human immu	780	5	16.7	232	11	Q99P51	Q99P51 mus musculu
708	5	16.7	214	15	Q79386	Q79386 human immu	781	5	16.7	232	16	Q8XIU6	Q8XIU6 clostridium
709	5	16.7	214	15	Q79387	Q79387 human immu	782	5	16.7	233	9	Q9ZXM2	Q9ZXM2 pseudomonas
710	5	16.7	214	15	Q79399	Q79399 human immu	783	5	16.7	233	9	Q9ZXM2	Q9ZXM2 bacterioph
711	5	16.7	214	15	Q79402	Q79402 human immu	784	5	16.7	234	16	P74626	P74626 synchocyst
712	5	16.7	214	15	Q79407	Q79407 human immu	785	5	16.7	235	2	Q9AGR7	Q9AGR7 streptococ
713	5	16.7	214	15	Q79412	Q79412 human immu	786	5	16.7	235	2	Q93MY4	Q93MY4 streptococ
714	5	16.7	214	15	Q79412	Q79412 human immu	787	5	16.7	236	10	Q9ZV06	Q9ZV06 arabidopsis
715	5	16.7	214	16	Q8YTB4	Q8YTB4 anabaena sp	788	5	16.7	236	16	Q9A108	Q9A108 streptococ
716	5	16.7	215	12	Q8QRY1	Q8QRY1 chimpanzee	789	5	16.7	237	5	Q20955	Q20955 caenorhabdi
717	5	16.7	215	15	Q77617	Q77617 human immu	790	5	16.7	237	10	Q9FHR2	Q9FHR2 arabidopsis
718	5	16.7	215	15	Q77619	Q77619 human immu	791	5	16.7	237	16	Q92G43	Q92G43 ticketsia
719	5	16.7	215	15	Q77621	Q77621 human immu	792	5	16.7	237	16	Q8YVC3	Q8YVC3 anabaena sp
720	5	16.7	215	15	Q77626	Q77626 human immu	793	5	16.7	238	2	Q9ZHV6	Q9ZHV6 nostoc punc
721	5	16.7	215	15	Q77655	Q77655 human immu	794	5	16.7	238	5	Q9WIM4	Q9WIM4 drosofila
722	5	16.7	215	15	Q77655	Q77655 human immu	795	5	16.7	238	16	Q8YV87	Q8YV87 anabaena sp
723	5	16.7	215	17	Q9YEL2	Q9YEL2 aetopyrum p	796	5	16.7	239	2	Q9RCL1	Q9RCL1 streptococ
724	5	16.7	217	15	Q97088	Q97088 human immu	797	5	16.7	239	15	Q8URQ9	Q8URQ9 human immu
725	5	16.7	217	15	Q97085	Q97085 human immu	798	5	16.7	239	15	Q8URQ8	Q8URQ8 human immu
726	5	16.7	217	15	Q8OAT8	Q8OAT8 human immu	799	5	16.7	239	15	Q8URQ7	Q8URQ7 human immu
727	5	16.7	218	16	Q91470	Q91470 pseudomonas	800	5	16.7	239	15	Q8URQ6	Q8URQ6 human immu
728	5	16.7	218	16	Q8ZNB3	Q8ZNB3 salmonella	801	5	16.7	239	15	Q8URQ5	Q8URQ5 human immu
729	5	16.7	218	16	Q8Z424	Q8Z424 salmonella	802	5	16.7	239	15	Q8URQ4	Q8URQ4 human immu
730	5	16.7	219	15	Q71074	Q71074 human immu	803	5	16.7	239	15	Q8URQ3	Q8URQ3 human immu
731	5	16.7	219	15	Q97087	Q97087 human immu	804	5	16.7	239	15	Q8URQ2	Q8URQ2 human immu
732	5	16.7	219	16	Q9JXY9	Q9JXY9 neisseria m	805	5	16.7	239	15	Q8URQ1	Q8URQ1 human immu
733	5	16.7	220	11	Q8VE86	Q8VE86 mus musculu	806	5	16.7	239	15	Q8URP9	Q8URP9 human immu
734	5	16.7	220	11	Q8VE86	Q8VE86 mus musculu	807	5	16.7	239	15	Q8URP8	Q8URP8 human immu
735	5	16.7	220	15	Q73287	Q73287 human immu	808	5	16.7	239	15	Q8URP7	Q8URP7 human immu
736	5	16.7	220	16	Q66346	Q66346 mycobacteri	809	5	16.7	239	15	Q8URP6	Q8URP6 human immu
737	5	16.7	220	17	Q9HJ46	Q9HJ46 thermoplasma	810	5	16.7	239	15	Q8URP5	Q8URP5 human immu
738	5	16.7	221	5	Q76328	Q76328 drosofila	811	5	16.7	239	15	Q8URP4	Q8URP4 human immu
739	5	16.7	221	10	Q9LFP1	Q9LFP1 arabidopsis	812	5	16.7	239	15	Q8URP3	Q8URP3 human immu
740	5	16.7	221	15	Q9YRX3	Q9YRX3 human immu	813	5	16.7	239	15	Q8URP2	Q8URP2 human immu
741	5	16.7	221	15	Q9YRX2	Q9YRX2 human immu	814	5	16.7	239	15	Q8URP1	Q8URP1 human immu
742	5	16.7	221	15	Q9YRX1	Q9YRX1 human immu	815	5	16.7	239	15	Q8URP0	Q8URP0 human immu
743	5	16.7	221	15	Q9YRX0	Q9YRX0 human immu	816	5	16.7	239	15	Q8URN9	Q8URN9 human immu
744	5	16.7	221	15	Q9YTB5	Q9YTB5 human immu	817	5	16.7	239	15	Q8URN8	Q8URN8 human immu
745	5	16.7	221	15	Q9YIX7	Q9YIX7 human immu	818	5	16.7	239	15	Q8URN7	Q8URN7 human immu
746	5	16.7	222	10	Q9LVV4	Q9LVV4 arabidopsis	819	5	16.7	239	15	Q8URN6	Q8URN6 human immu

820	5	16.7	239	15	Q8URN5	Q8urn5 human	immun	893	5	16.7	243	15	Q9E020	Q9e020 human	immun
821	5	16.7	239	15	Q8URN4	Q8urn4 human	immun	894	5	16.7	243	15	Q9E019	Q9e019 human	immun
822	5	16.7	239	15	Q8URN3	Q8urn3 human	immun	895	5	16.7	243	15	Q9E018	Q9e018 human	immun
823	5	16.7	239	15	Q8URN2	Q8urn2 human	immun	896	5	16.7	243	15	Q9E017	Q9e017 human	immun
824	5	16.7	239	15	Q8URN1	Q8urn1 human	immun	897	5	16.7	243	15	Q9E016	Q9e016 human	immun
825	5	16.7	239	15	Q8URN0	Q8urn0 human	immun	898	5	16.7	243	15	Q9E015	Q9e015 human	immun
826	5	16.7	239	15	Q8URN9	Q8urn9 human	immun	899	5	16.7	243	15	Q9E014	Q9e014 human	immun
827	5	16.7	239	15	Q8URN8	Q8urn8 human	immun	900	5	16.7	243	15	Q9E013	Q9e013 human	immun
828	5	16.7	239	15	Q8URN7	Q8urn7 human	immun	901	5	16.7	243	15	Q9E012	Q9e012 human	immun
829	5	16.7	239	15	Q8URN6	Q8urn6 human	immun	902	5	16.7	243	15	Q9E011	Q9e011 human	immun
830	5	16.7	239	16	Q9PGU0	Q9pgu0 xylella fas	903	5	16.7	243	15	Q9E010	Q9e010 human	immun	
831	5	16.7	239	16	Q9KKR5	Q9kkrs5 vibrio chol	904	5	16.7	243	15	Q9E009	Q9e009 human	immun	
832	5	16.7	239	16	Q97MW9	Q97mw9 clostridium	905	5	16.7	243	15	Q9E008	Q9e008 human	immun	
833	5	16.7	240	11	Q9CUL6	Q9cul6 mus musculi	906	5	16.7	243	16	Q8X7Y5	Q8x7y5 escherichia		
834	5	16.7	241	5	Q9XWMS	Q9xwms caenorhabdi	907	5	16.7	243	17	Q9H1F1	Q9h1f1 thermoplasma		
835	5	16.7	241	5	001464	001464 caenorhabdi	908	5	16.7	244	4	013784	013784 homo sapien		
836	5	16.7	241	12	011309	011309 molluscum c	909	5	16.7	244	5	020804	020804 caenorhabdi		
837	5	16.7	242	2	Q9ZA27	Q9za27 streptomyce	910	5	16.7	245	2	Q9X7K5	Q9x7k5 rhizobium g		
838	5	16.7	242	2	050391	050391 mycobacteri	911	5	16.7	245	4	Q9NYC4	Q9nyc4 homo sapien		
839	5	16.7	242	10	Q91LGB9	Q91lgb9 oryza sativ	912	5	16.7	245	16	Q67870	Q67870 aquifex aeo		
840	5	16.7	242	16	Q9BJ89	Q9bj89 rhizobium 1	913	5	16.7	246	2	Q9EUM7	Q9eum7 listeria in		
841	5	16.7	243	10	Q919U2	Q919u2 lagopus lag	914	5	16.7	246	11	061435	061435 mus musculi		
842	5	16.7	243	15	Q9E072	Q9e072 human	immun	915	5	16.7	246	16	Q8YTC1	Q8ytc1 anabaena sp	
843	5	16.7	243	15	Q9E071	Q9e071 human	immun	916	5	16.7	247	2	032546	032546 escherichia	
844	5	16.7	243	15	Q9E070	Q9e070 human	immun	917	5	16.7	247	10	Q9F1X5	Q9f1x5 arabidopsis	
845	5	16.7	243	15	Q9E069	Q9e069 human	immun	918	5	16.7	247	16	Q8Z1Z8	Q8z1z8 salmonella	
846	5	16.7	243	15	Q9E068	Q9e068 human	immun	919	5	16.7	247	16	Q8Z2J5	Q8z2j5 salmonella	
847	5	16.7	243	15	Q9E067	Q9e067 human	immun	920	5	16.7	247	16	Q9ZRA8	Q9zra8 rhizobium m	
848	5	16.7	243	15	Q9E066	Q9e066 human	immun	921	5	16.7	248	2	Q9AGT0	Q9agt0 staphylococ	
849	5	16.7	243	15	Q9E065	Q9e065 human	immun	922	5	16.7	248	5	Q25676	Q25676 plasmodium	
850	5	16.7	243	15	Q9E064	Q9e064 human	immun	923	5	16.7	248	16	Q99W72	Q99w72 staphylococ	
851	5	16.7	243	15	Q9E063	Q9e063 human	immun	924	5	16.7	248	16	Q8UGK3	Q8ugk3 agrobacteri	
852	5	16.7	243	15	Q9E062	Q9e062 human	immun	925	5	16.7	248	17	Q26833	Q26833 methanobact	
853	5	16.7	243	15	Q9E061	Q9e061 human	immun	926	5	16.7	249	10	Q8W5E4	Q8w5e4 oryza sativ	
854	5	16.7	243	15	Q9E060	Q9e060 human	immun	927	5	16.7	249	16	Q9JGM5	Q9jgm5 salmonella	
855	5	16.7	243	15	Q9E059	Q9e059 human	immun	928	5	16.7	250	10	Q9LE11	Q9le11 hordeum vul	
856	5	16.7	243	15	Q9E058	Q9e058 human	immun	929	5	16.7	250	10	Q9ASV8	Q9asv8 arabidopsis	
857	5	16.7	243	15	Q9E057	Q9e057 human	immun	930	5	16.7	251	2	Q9APS6	Q9aps6 uncultured	
858	5	16.7	243	15	Q9E056	Q9e056 human	immun	931	5	16.7	251	2	Q9APN9	Q9apn9 uncultured	
859	5	16.7	243	15	Q9E055	Q9e055 human	immun	932	5	16.7	251	16	Q8ZC06	Q8zcg6 salmonella	
860	5	16.7	243	15	Q9E054	Q9e054 human	immun	933	5	16.7	251	16	Q8Z891	Q8z8z1 salmonella	
861	5	16.7	243	15	Q9E053	Q9e053 human	immun	934	5	16.7	251	16	Q8X8Z2	Q8x8z2 escherichia	
862	5	16.7	243	15	Q9E051	Q9e051 human	immun	935	5	16.7	252	5	Q9NJG8	Q9njg8 caenorhabdi	
863	5	16.7	243	15	Q9E050	Q9e050 human	immun	936	5	16.7	252	10	Q94167	Q94167 potentilla	
864	5	16.7	243	15	Q9E049	Q9e049 human	immun	937	5	16.7	252	16	Q8YEW2	Q8yew2 bruceella me	
865	5	16.7	243	15	Q9E048	Q9e048 human	immun	938	5	16.7	253	2	Q9WTC0	Q9wtc0 escherichia	
866	5	16.7	243	15	Q9E047	Q9e047 human	immun	939	5	16.7	253	2	Q937T4	Q937t4 acinetobact	
867	5	16.7	243	15	Q9E046	Q9e046 human	immun	940	5	16.7	254	5	Q9N5X6	Q9n5x6 caenorhabdi	
868	5	16.7	243	15	Q9E045	Q9e045 human	immun	941	5	16.7	254	17	059523	059523 pyrococcus	
869	5	16.7	243	15	Q9E044	Q9e044 human	immun	942	5	16.7	255	10	Q9ZWM8	Q9zwm8 oryza sativ	
870	5	16.7	243	15	Q9E043	Q9e043 human	immun	943	5	16.7	255	13	Q91911	Q91911 oncorhynch	
871	5	16.7	243	15	Q9E042	Q9e042 human	immun	944	5	16.7	256	10	Q963Z6	Q963z6 arabidopsis	
872	5	16.7	243	15	Q9E041	Q9e041 human	immun	945	5	16.7	256	12	Q9UG85	Q9ug85 ct virus. o	
873	5	16.7	243	15	Q9E040	Q9e040 human	immun	946	5	16.7	257	3	Q943Z2	Q943z2 schizosacch	
874	5	16.7	243	15	Q9E039	Q9e039 human	immun	947	5	16.7	257	16	Q9Z2AN0	Q9z2an0 listeria in	
875	5	16.7	243	15	Q9E038	Q9e038 human	immun	948	5	16.7	257	16	Q8YEB3	Q8yeb3 listeria mo	
876	5	16.7	243	15	Q9E037	Q9e037 human	immun	949	5	16.7	258	10	Q80714	Q80714 arabidopsis	
877	5	16.7	243	15	Q9E036	Q9e036 human	immun	950	5	16.7	258	16	Q8ZRB4	Q8zrb4 salmonella	
878	5	16.7	243	15	Q9E035	Q9e035 human	immun	951	5	16.7	258	16	Q8Z8R0	Q8z8r0 salmonella	
879	5	16.7	243	15	Q9E034	Q9e034 human	immun	952	5	16.7	259	2	Q9S3U5	Q9s3u5 bacteroides	
880	5	16.7	243	15	Q9E033	Q9e033 human	immun	953	5	16.7	259	10	Q9SNO0	Q9sno0 oryza sativ	
881	5	16.7	243	15	Q9E032	Q9e032 human	immun	954	5	16.7	259	17	Q8TXZ2	Q8txz2 pyrococcus	
882	5	16.7	243	15	Q9E031	Q9e031 human	immun	955	5	16.7	260	4	Q8TCH6	Q8tch6 homo sapien	
883	5	16.7	243	15	Q9E030	Q9e030 human	immun	956	5	16.7	260	12	Q8UZ53	Q8uz53 indian cass	
884	5	16.7	243	15	Q9E029	Q9e029 human	immun	957	5	16.7	260	16	Q8YLU2	Q8ylu2 anabaena sp	
885	5	16.7	243	15	Q9E028	Q9e028 human	immun	958	5	16.7	261	4	Q9BSG5	Q9bsg5 homo sapien	
886	5	16.7	243	15	Q9E027	Q9e027 human	immun	959	5	16.7	261	16	Q98CP4	Q98cp4 rhizobium 1	
887	5	16.7	243	15	Q9E026	Q9e026 human	immun	960	5	16.7	262	16	Q8Y153	Q8y153 anabaena sp	
888	5	16.7	243	15	Q9E025	Q9e025 human	immun	961	5	16.7	263	4	Q96MW1	Q96mw1 homo sapien	
889	5	16.7	243	15	Q9E024	Q9e024 human	immun	962	5	16.7	263	15	Q9WDN7	Q9wdn7 human	immun
890	5	16.7	243	15	Q9E023	Q9e023 human	immun	963	5	16.7	263	15	Q9WDU3	Q9wdu3 human	immun
891	5	16.7	243	15	Q9E022	Q9e022 human	immun	964	5	16.7	263	15	Q9WDU5	Q9wdu5 human	immun
892	5	16.7	243	15	Q9E021	Q9e021 human	immun	965	5	16.7	263	15	Q9WDU8	Q9wdu8 human	immun

966	5	16.7	263	15	Q9WDV0	Q9wdv0 human immun
967	5	16.7	263	15	Q9WDV6	Q9wdv6 human immun
968	5	16.7	263	15	Q9WDV8	Q9wdv8 human immun
969	5	16.7	263	15	Q9WDW0	Q9wdw0 human immun
970	5	16.7	263	15	Q9WDW3	Q9wdw3 human immun
971	5	16.7	263	15	Q9WDW5	Q9wdw5 human immun
972	5	16.7	263	15	Q9WDW7	Q9wdw7 human immun
973	5	16.7	263	15	Q9WDW8	Q9wdw8 human immun
974	5	16.7	263	15	Q9WDW9	Q9wdw9 human immun
975	5	16.7	263	15	Q9WDX0	Q9wdx0 human immun
976	5	16.7	263	15	Q9WDX5	Q9wdx5 human immun
977	5	16.7	263	15	Q9WDX6	Q9wdx6 human immun
978	5	16.7	263	15	Q9WE18	Q9we18 human immun
979	5	16.7	263	15	Q9WE19	Q9we19 human immun
980	5	16.7	263	15	Q9WE10	Q9we10 human immun
981	5	16.7	263	15	Q9WEJ1	Q9wej1 human immun
982	5	16.7	263	15	Q9WEJ2	Q9wej2 human immun
983	5	16.7	263	15	Q9WEJ4	Q9wej4 human immun
984	5	16.7	263	15	Q9WEJ5	Q9wej5 human immun
985	5	16.7	263	15	Q9WEJ7	Q9wej7 human immun
986	5	16.7	263	15	Q9WEJ8	Q9wej8 human immun
987	5	16.7	263	15	Q9WEJ9	Q9wej9 human immun
988	5	16.7	263	15	Q9WEJ3	Q9wej3 human immun
989	5	16.7	263	15	Q9WE17	Q9we17 human immun
990	5	16.7	263	15	Q9WEM0	Q9wem0 human immun
991	5	16.7	263	15	Q9WEM1	Q9wem1 human immun
992	5	16.7	263	15	Q9WEM4	Q9wem4 human immun
993	5	16.7	263	15	Q9WEM7	Q9wem7 human immun
994	5	16.7	263	15	Q9WEM8	Q9wem8 human immun
995	5	16.7	263	16	Q8YK11	Q8yK11 arabidena sp
996	5	16.7	264	4	Q9H9Q2	Q9h9q2 homo sapien
997	5	16.7	264	4	Q96BZ8	Q96bz8 homo sapien
998	5	16.7	264	11	O88547	O88547 mus musculu
999	5	16.7	264	11	Q92JG4	Q92jg4 mus musculu
1000	5	16.7	264	15	Q9WDM2	Q9wdm2 human immun

## ALIGNMENTS

RESULT 1

Q9F9R1 PRELIMINARY; PRT; 207 AA.

AC Q9F9R1; 01-MAR-2001 (TREMBlrel. 16, Created)

DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE Superoxide dismutase.

EN SOD.

OS Mycobacterium paratuberculosis.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.

OX NCBI\_TaxID=1770;

XX [1]

RN SEQUENCE FROM N.A.

RP STRAIN=K-10;

RC Liu X., Peng Z., Carillo J., Barletta R.G.;

RA "Mycobacterium paratuberculosis manganese superoxide dismutase.";

RT Submitted (AUG-1999) to the EMBL/Genbank/DBJ databases.

CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE

CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).

CC -1- CATALYTIC ACTIVITY: 2 SUPEROXIDE + 2 H(+) = O(2) + H(2)O(2).

CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE

CC FAMILY.

DR EMBL; AF180816; AAG09425.1; -.

DR HSSP; P17670; IIDS.

DR InterPro; IPR001189; SODismutase.

DR Pfam; PF00081; sodfe\_1.

DR Pfam; PF02777; sodfe\_C1.

DR ProDom; PD000475; SODismutase; 1.

DR PROSITE; PS00088; SOD\_MN; 1.

KW Oxidoreductase.

SO SEQUENCE 207 AA; 23030 MM; EDA8C2EB40ED428D CRC64;

Query Match 26.7%; Score 8; DB 2; Length 207;

Best Local Similarity 100.0%; Pred. No. 0.68;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 ISGQINEI 25

Db 20 ISGQINEI 27

## RESULT 2

Q9AM00 PRELIMINARY; PRT; 210 AA.

AC Q9AM00; 01-JUN-2001 (TREMBlrel. 17, Created)

DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE Superoxide dismutase.

OS Mycobacterium paratuberculosis.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.

OX NCBI\_TaxID=1770;

XX [1]

RN SEQUENCE FROM N.A.

RP Dheendhayalan V., Chang Y.F.;

RA "Mycobacterium avium subsp. paratuberculosis superoxide dismutase

RT gene.";

RL Submitted (JAN-2001) to the EMBL/Genbank/DBJ databases.

CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE

CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).

CC -1- CATALYTIC ACTIVITY: 2 SUPEROXIDE + 2 H(+) = O(2) + H(2)O(2).

CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE

CC FAMILY.

DR EMBL; AF33434; AAG50084.1; -.

DR HSSP; P17670; IIDS.

DR InterPro; IPR001189; SODismutase.

DR Pfam; PF00081; sodfe\_1.

DR Pfam; PF02777; sodfe\_C1.

DR ProDom; PD000475; SODismutase; 1.

DR PROSITE; PS00088; SOD\_MN; 1.

KW Oxidoreductase.

SO SEQUENCE 210 AA; 23357 MM; E492E31352B2E20 CRC64;

Query Match 26.7%; Score 8; DB 2; Length 210;

Best Local Similarity 100.0%; Pred. No. 0.69;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 ISGQINEI 25

Db 20 ISGQINEI 27

RESULT 3

P79022 PRELIMINARY; PRT; 227 AA.

AC P79022; 01-MAY-1997 (TREMBlrel. 03, Created)

DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE Superoxide dismutase (EC 1.15.1.1) precursor.

EN MNSOD.

OS Candida sp. Fung95.

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; microspor Saccharomycetales; Candida.

OX NCBI\_TaxID=159257;

XX [1]

RN SEQUENCE FROM N.A.

RP Choi S.-Y., Nam Y.-S., Hong Y.M.;

RA "Molecular cloning and characterization of Mn-superoxide dismutase

RT gene from Candida sp.";

RL Korean J. Microbiol. 35:309-314(1997).

CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE

CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).

CC -1- CATALYTIC ACTIVITY: 2 SUPEROXIDE + 2 H(+) = O(2) + H(2)O(2).  
 CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE  
 CC FAMILY.  
 DR EMBL: Y11598; CAA72335.1; -.  
 DR HSSP: P04179; IAP6.  
 DR InterPro: IPR001189; SODismutase.  
 DR Pfam: PF00081; sode, 1.  
 DR Pfam: PF02777; sode C, 1.  
 DR ProDom: PD000475; SODismutase; 1.  
 DR PROSITE: PS00088; SOD MN; 1.  
 KM Oxidoreductase; Signal.  
 FT SIGNAL 1 30  
 FT CHAIN 1 227 SUPEROXIDE DISMUTASE.  
 FT SEQUENCE 227 AA; 25227 MW; 29E7F6E3C7743DB0 CRC64;  
 SQ  
 Query Match 26.7%; Score 8; DB 3; Length 227;  
 Best Local Similarity 100.0%; Pred. No. 0.74;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 18 ISGQINEI 25  
 |||||  
 Db 46 ISGQINEI 53  
 RESULT 4  
 ID Q93X78 PRELIMINARY; PRT; 226 AA.  
 AC Q93X78;  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Beta-glucosidase (EC 3.2.1.21) (Fragment).  
 OS Vitis vinifera (Grape).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Vitaceae;  
 OC Vitif.  
 OX NCBI\_TaxID=29760;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. SHIRAZ; TISSUE=FRUIT;  
 RA Sarry J.-E., Romieu C., Terrier N.;  
 RT "Beta-glucoside hydrolase from grape berries."  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AY039034; AAK72100.1; -.  
 DR InterPro: IPR001360; GH 1.  
 DR Pfam: PF00232; Glyco\_hydro\_1; 1.  
 DR ProDom: PD000650; GH\_1; 1.  
 DR Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC Non TER 1  
 FT NON TER 1  
 FT SEQUENCE 226 AA; 25610 MW; 2153960F27E61432 CRC64;  
 SQ

Query Match 23.3%; Score 7; DB 10; Length 226;  
 Best Local Similarity 100.0%; Pred. No. 8.6;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 12 SATEPYI 18  
 |||||  
 Db 83 SATEPYI 89

RESULT 5  
 ID Q9CSH7 PRELIMINARY; PRT; 312 AA.  
 AC Q9CSH7;  
 DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Putative beta-glucosidase protein.  
 GN MRC8.6/AT3G18080.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC euroside II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,  
 RA Goldsmith A.D., Lee J.M., Quach H.L., Tortum M., Yu G., Bowser L.,  
 RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,  
 RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,  
 RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,  
 RA Seki M., Shim P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,  
 RA Theologis A.;  
 RT "Full Length cDNA of gene MRC8.6/AT3G18080 (GI:9294063).";  
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,  
 RA Dale J.M., Gibson H.A., Goldsmith A.D., Jiang P.X., Lee J.M.,  
 RA Quach H.L., Tang C., Tortum M., Yu G., Bowser L., Carninci P.,  
 RA Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T., Kamiya A.,  
 RA Karlin-Neumann G., Kawai J., Kim C., Koesema E., Lam B., Lin J.,  
 RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,  
 RA Sakurai T., Satou M., Seki M., Shim P., Southwick A., Tracy S.E.,  
 RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;  
 RT "Full Length cDNA of gene MRC8.6/AT3G18080 (GI:9294063).";  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF360240; AAK25950.1; -.  
 DR EMBL: AY040038; AAK64096.1; -.  
 DR HSSP: P26205; ICBG.  
 DR InterPro: IPR001360; GH 1.  
 DR Pfam: PF00232; Glyco\_hydro\_1; 1.  
 DR PRINTS: PR00131; GHYDLASE1.  
 DR ProDom: PD000650; GH 1; 1.  
 SQ SEQUENCE 312 AA; 36162 MW; 908D8E22553489F CRC64;  
 SQ

Query Match 23.3%; Score 7; DB 10; Length 312;  
 Best Local Similarity 100.0%; Pred. No. 11;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 12 SATEPYI 18  
 |||||  
 Db 36 SATEPYI 42

RESULT 6  
 ID Q942J1 PRELIMINARY; PRT; 462 AA.  
 AC Q942J1;  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Putative beta-glucosidase.  
 GN P045H0310.1.  
 OS Oryza sativa (Rice).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Eriarthoideae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. NIPPONBARE;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare(GAS) genomic DNA, chromosome 1, PAC  
 clone: P045H0310.1.";  
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AP003435; BAB68015.1; -.  
 DR InterPro: IPR001360; GH 1.  
 DR Pfam: PF00232; Glyco\_hydro\_1; 1.  
 DR ProDom: PD000650; GH 1; 1.  
 SQ SEQUENCE 462 AA; 53007 MW; A4FBA6A625AF052 CRC64;  
 SQ

Query Match 23.3%; Score 7; DB 10; Length 462;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 SATEPYI 18  
 |||||  
 Db 153 SATEPYI 159

## RESULT 7

Q9LV34 PRELIMINARY; PRT; 495 AA.

AC Q9LV34  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Beta-glucosidase.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsiis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=COLUMBIA;  
 RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;  
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
 RT  
 RT TAC and BAC clones."  
 RL DNA Res. 7:217-221(2000).  
 DR EMBL; AB020749; BAB02019.1; -.  
 DR HSSP; P26205; ICBG.  
 DR InterPro; IPR001360; GH\_1.  
 DR Pfam; PF00232; Glyco\_hydro\_1; 1.  
 DR PRINTS; PR00131; GLHYDRLASE1.  
 DR PRODOM; PD000650; GH\_1; 1.  
 SQ SEQUENCE 495 AA; 56496 MW; 85D92F613B79C22 CRC64;

Query Match 23.3%; Score 7; DB 10; Length 495;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 12 SATEPYI 18  
 |||||  
 Db 219 SATEPYI 225

## RESULT 8

Q42975 PRELIMINARY; PRT; 504 AA.

AC Q42975  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Beta-glucosidase (EC 3.2.1.21).  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Euharicotidae; Oryzae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE OF 1-123 FROM N.A.  
 RC STRAIN=NIPPONBARE; TISSUE=ROOT;  
 RA Yuzo M., Takuji S.;  
 RT "Oryza sativa beta-glucosidase mRNA sequence."  
 RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NIPPONBARE; TISSUE=ROOT;  
 RA Eesen A., Opasitiri R., Ketudat Cairns J., Akiyama T.;

RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U28047; AAA84906.2; -.  
 KW Hydrolase; Glycosidase.  
 SQ SEQUENCE 504 AA; 56906 MW; 5D93BF0605E43257 CRC64;

Query Match 23.3%; Score 7; DB 10; Length 504;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 SATEPYI 18  
 |||||  
 Db 232 SATEPYI 238

## RESULT 9

Q24524 PRELIMINARY; PRT; 507 AA.

AC Q24524  
 DT 01-JAN-1998 (TREMBlrel. 05, Created)  
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Linamarase (EC 3.2.1.21).  
 GN PLIN-GEN.  
 OS Manihot esculenta (Cassava) (Manioc).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Malpighiales; Euphorbiaceae; Manihot.  
 OX NCBI\_TaxID=3983;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC McMahon J.M., Sayre R.T.;  
 RT "Genomic sequence for a linamarase gene from cassava (Manihot  
 esculenta Crantz)."  
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U95298; AAB71381.1; -.  
 DR HSSP; P26205; ICBG.  
 DR InterPro; IPR001360; GH\_1.  
 DR Pfam; PF00232; Glyco\_hydro\_1; 1.  
 DR PRINTS; PR00131; GLHYDRLASE1.  
 DR PRODOM; PD000650; GH\_1; 1.  
 KW Glycosidase; Hydrolase.  
 SQ SEQUENCE 507 AA; 59324 MW; 9CF33AD8755A7AF CRC64;

Query Match 23.3%; Score 7; DB 10; Length 507;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 SATEPYI 18  
 |||||  
 Db 206 SATEPYI 212

## RESULT 10

Q9LV33 PRELIMINARY; PRT; 512 AA.

AC Q9LV33  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Beta-glucosidase.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsiis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=COLUMBIA;  
 RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;  
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=COLUMBIA;



```

RX MEDLINE=20363099; PubMed=10907853;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT Sequence features of the regions of 4,251,695 bp covered by ninety P1,
RT TAC and BAC clones.";
RL DNA Res. 7:217-221(2000).
DR EMBL; AB020749; BAB02020.1; -.
DR HSSP; P26205; ICBG.
DR InterPro; IPR001360; GH 1.
DR Pfam; PF00232; Glyco_hydro_1; 1.
DR PRINTS; PR00131; GLHYDRLASE1.
DR PRODOM; PD000650; GH 1; 1.
SQ SEQUENCE 512 AA; 58983 MW; A870F9BB20C342E7 CRC64;

Query Match 23.3%; Score 7; DB 10; Length 512;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 SATEPYI 18
    |||||
    236 SATEPYI 242

RESULT 11
O8VWL8 PRELIMINARY; PRT; 514 AA.
ID O8VWL8;
AC O8VWL8;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Beta-mannosidase (Beta-mannosidase enzyme).
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; eusteridae I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=CV; GLAMOUR;
RA Mo B.; Bewley D.;
RT "Beta-mannosidase gene from tomato (Lewside 2).";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV; GLAMOUR; TISSUE=SEED;
RA Mo B.; Bewley J.;
RT "Lycopersicon esculentum beta-mannosidase (Lewside) gene encoding
RT beta-mannosidase enzyme, complete sequence.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF413204; ALJ37719.1; -.
DR EMBL; AF413204; ALJ37714.1; -.
DR InterPro; IPR001360; GH 1.
DR Pfam; PF00232; Glyco_hydro_1; 1.
DR PRINTS; PR00131; GLHYDRLASE1.
DR PRODOM; PD000650; GH 1; 1.
DR PROSITE; PS00653; GLYCOSYL_HYDROL_F1_2; UNKNOWN 1.
SQ SEQUENCE 514 AA; 59288 MW; 0A3855084EDA1971 CRC64;

Query Match 23.3%; Score 7; DB 10; Length 514;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 SATEPYI 18
    |||||
    239 SATEPYI 245

RESULT 12
O41172 PRELIMINARY; PRT; 531 AA.
ID O41172;
AC O41172;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

```

```

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Linamarase.
GN LINAMARASE.
OS Manihot esculenta (Cassava) (Manioc).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucots I; Malpighiales; Euphorbiaceae; Manihot.
OX NCBI_TaxID=3983;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92264724; PubMed=1586156;
RA Hughes M.A., Brown K., Pancoro A., Murray B.S., Oxtoby E., Hughes J.;
RT "A molecular and biochemical analysis of the structure of the
RT cyanogenic beta-glucosidase (linamarase) from cassava (Manihot
RT esculenta Cranz).";
RL Arch. Biochem. Biophys. 295:273-279(1992).
DR EMBL; S35175; AB22162.1; -.
DR HSSP; P26205; ICBG.
DR InterPro; IPR001360; GH 1.
DR Pfam; PF00232; Glyco_hydro_1; 1.
DR PRINTS; PR00131; GLHYDRLASE1.
DR PRODOM; PD000650; GH 1; 1.
SQ SEQUENCE 531 AA; 61373 MW; 4FE14D39FBA0F17 CRC64;

Query Match 23.3%; Score 7; DB 10; Length 531;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 SATEPYI 18
    |||||
    230 SATEPYI 236

RESULT 13
O40283 PRELIMINARY; PRT; 541 AA.
ID O40283;
AC O40283;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Beta glucosidase precursor.
GN BGLA.
OS Manihot esculenta (Cassava) (Manioc).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucots I; Malpighiales; Euphorbiaceae; Manihot.
OX NCBI_TaxID=3983;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CBG3;
RL Little S.; Keresztesy Z., Hughes J., Hughes M.A.;
RT "A genomic cyanogenic beta-glucosidase gene from Cassava (Accession
RT No. X94986). (PGR98-148).";
RL Plant Physiol. 117:1526-1526(1998).
DR EMBL; X94986; CA64442.1; -.
DR HSSP; P26205; ICBG.
DR InterPro; IPR001360; GH 1.
DR Pfam; PF00232; Glyco_hydro_1; 1.
DR PRINTS; PR00131; GLHYDRLASE1.
DR PRODOM; PD000650; GH 1; 1.
DR PROSITE; PS00653; GLYCOSYL_HYDROL_F1_2; 1.
KW SIGNAL.
FT SIGNAL. 1 28 POTENTIAL.
FT CHAIN 29 541 BETA GLUCOSIDASE.
SQ SEQUENCE 541 AA; 63095 MW; 01735B5D4429EDA CRC64;

Query Match 23.3%; Score 7; DB 10; Length 541;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 SATEPYI 18
    |||||
    241 SATEPYI 247

```

```
RESULT 14
Q8SVR3 PRELIMINARY; PRT; 620 AA.
ID Q8SVR3;
AC Q8SVR3;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
GN Hypothetical protein ECU04_1410.
OS ECU04_1410.
OC Eukaryota; Microsporidia; Unikaryonidae; Encephalitozoon.
OX NCBI_TaxID=6035;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GB-M1;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
SQ SEQUENCE FROM N.A.
STRAIN=GB-M1;
MEDLINE=21576510; PubMed=11719806;
RA Katinka M.D., Duprat S., Cornillot E., Metenier G., Thomarat P.,
RA Prenster G., Barthe V., Peyretailade E., Brotier P., Wincker P.,
RA Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,
RA Weissbach J., Vivares C.P.;
RT "Genome sequence and gene compaction of the eukaryote parasite
RT Encephalitozoon cuniculi."
RL Nature 414:450-453(2001).
DR EMBL; AL590444; CAD25330.1; -.
KW Hypothetical protein.
SQ SEQUENCE 620 AA; 69120 MW; CED63EACFE52191 CRC64;

Query Match 23.3%; Score 7; DB 5; Length 620;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KYSLPEL 7
Db 264 KYSLPEL 270

RESULT 15
Q9U275 PRELIMINARY; PRT; 807 AA.
ID Q9U275;
AC Q9U275;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
GN Y50E8A.16 protein.
OS Y50E8A.16.
OC Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodermidae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Steward C.A.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
SQ SEQUENCE FROM N.A.
MEDLINE=99069613; PubMed=9851916;
none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology."
RL Science 282:2012-2018(1998).
CC -1 SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR EMBL; AL117200; CAB60586.1; -.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001140; ABCtransportTM.
DR InterPro; IPR003439; ABC_transport.
DR Pfam; PF00664; ABC_membrane; 1.
```

```
DR Pfam; PF00005; ABC_tran; 1.
DR ProDom; PD000006; ABC_transport; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW ATP-binding; Transport.
SQ SEQUENCE 807 AA; 91163 MW; B9ABC598966DF1EF CRC64;

Query Match 23.3%; Score 7; DB 5; Length 807;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PELDYEF 11
Db 532 PELDYEF 538

RESULT 16
Q9AMT5 PRELIMINARY; PRT; 936 AA.
ID Q9AMT5;
AC Q9AMT5;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
GN ID880.
OS Bradyrhizobium japonicum.
OC Bradyrhizobium japonicum; alpha subdivision; Rhizobiaceae group;
OC Bradyrhizobium group; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=110SPC4;
RX MEDLINE=21101824; PubMed=11157954;
RA Gottfert M., Rothlisberger S., Kundig C., Beck C., Marty R.,
RA Hennecke H.;
RT "Potential symbiosis-specific genes uncovered by sequencing a 410-kb
RT DNA region of the Bradyrhizobium japonicum chromosome."
RL J. Bacteriol. 183:1405-1412(2001).
DR EMBL; AF322013; AAG61062.1; -.
SQ SEQUENCE 936 AA; 104325 MW; 9AD94C465044F904 CRC64;

Query Match 23.3%; Score 7; DB 2; Length 936;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 EFSATEP 16
Db 393 EFSATEP 399

RESULT 17
Q94QB2 PRELIMINARY; PRT; 1011 AA.
ID Q94QB2;
AC Q94QB2;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
GN Phosphoenolpyruvate carboxylase (EC 4.1.1.31).
OS SVR6C.
OS Synecococcus vulcanus (Thermosynechococcus vulcanus).
OG Nucleomorph.
OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.
OX NCBI_TaxID=32053;
RN [1]
RP SEQUENCE FROM N.A.
RA Chen L.M., Omiya T., Hata S., Inoue Y., Imai K.;
RT "Molecular characterization of Synecococcus vulcanus
RT phosphoenolpyruvate carboxylase."
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB057454; BAB64533.1; -.
DR InterPro; IPR001449; PEPcase.
DR Pfam; PF00311; PEPcase; 1.
DR PROSITE; PS00393; PEPcase_2; UNKNOWN_1.
```



RT "Genetic organization of the mobilization region of the plasmid pHE1  
 RL from *Halomonas elongata*.";  
 RL Syst. Appl. Microbiol. 22:520-529(1999).  
 DR EMBL; AJ243735; CAB96960.1; -.  
 KW Plasmid.  
 SO SEQUENCE 150 AA; 16460 MW; 77B7691EF76DBE93 CRC64;

Query Match 20.0%; Score 6; DB 2; Length 150;  
 Best Local Similarity 100.0%; Pred. No. 70;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 20 GQINEI 25  
 |||||  
 Db 94 GQINEI 99

## RESULT 22

O9CTJ9 PRELIMINARY; PRT; 153 AA.

O9CTJ9;  
 01-JUN-2001 (TREMBlrel. 17, Created)  
 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE 18 days embryo cDNA, RIKEN full-length enriched library,  
 DE clone:110002121, full insert sequence (fragment).  
 GN 2210412K09RIK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=EMBRYO;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kaubakawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuhl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
 RA Schirral L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blakes J., Boftelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Guelinich S., Hill D., Hofmann K., Hume D.A., Kamuya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombere P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Saeki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmink L.,  
 RA Wyshaw-Boris A., Yoshida K., Haegawa Y., Kawaji H., Kohlsaki S.,  
 RA Hayashizaki Y.;  
 RA "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 DR EMBL; AK003313; BAB2709.1; -.  
 DR MGD; MGI:1924209; 2210412K09RIK.  
 FT NON TER 1  
 SO SEQUENCE 153 AA; 16332 MW; E780CE2347A4C3B2 CRC64;

Query Match 20.0%; Score 6; DB 11; Length 153;  
 Best Local Similarity 100.0%; Pred. No. 72;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 SLPELD 8  
 |||||  
 Db 59 SLPELD 64

## RESULT 23

O9BPM2 PRELIMINARY; PRT; 156 AA.

ID O9BPM2  
 AC O9BPM2;  
 DT 01-JUN-2001 (TREMBlrel. 17, Created)

DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Cysteine proteinase (Fragment).  
 GN Cpl.  
 OS Metagonimus yokogawai.  
 OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;  
 OC Opisthorchiida; Opisthorchiata; Opisthorchiidae; Heterophyidae;  
 OC Metagonimus.  
 OX NCBI\_TaxID=84529;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RA Park H., Yun H.C., Kim K.Y., Park S.Y., Park S.K.;  
 RT "Metagonimus yokogawai cysteine proteinase (Cpl) mRNA."  
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF121840; AAC52658.1; -.  
 DR HSSP; P07711; 1CS8.  
 DR MEROPS; C01.032; -.  
 DR InterPro; IPR000668; Peptidase\_C1.  
 DR InterPro; IPR000169; SHprot\_acsite.  
 DR InterPro; IPR000834; Zn\_cardopept.  
 DR Pfam; PF00112; Peptidase\_C1; 1.  
 DR ProDom; PD000158; Peptidase\_C1; 1.  
 DR PROSITE; PS00133; CARBOXYPEPT\_ZN\_2; UNKNOWN\_1.  
 DR PROSITE; PS00639; THIOL\_PROTEASE\_HIS; 1.  
 FT NON TER 1  
 FT NON TER 1  
 SO SEQUENCE 156 AA; 16938 MW; 26ED609F5F315B15 CRC64;

Query Match 20.0%; Score 6; DB 5; Length 156;  
 Best Local Similarity 100.0%; Pred. No. 73;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 16 PYISQ 21  
 |||||  
 Db 63 PYISQ 68

## RESULT 24

O99ZV3 PRELIMINARY; PRT; 163 AA.

DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE Purative phosphotransferase system (PTS), enzyme II component B.  
 GN SPY1058.  
 OS Streptococcus pyogenes.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;  
 OC Streptococcaceae; Streptococcus.  
 OX NCBI\_TaxID=1314;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SF370 / ATCC 700294 / SEROTYPE M1;  
 RX MEDLINE=21192684; PubMed=11296296;  
 RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,  
 RA Primeaux C., Sezate S., Suvorov A.N., Kenyon S., Lai H.S., Lin S.P.,  
 RA Qian Y., Jia H.G., Najjar F.Z., Ren O., Zhu H., Song L., White J.,  
 RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;  
 RT "Complete genome sequence of an M1 strain of *Streptococcus pyogenes*."  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).  
 DR EMBL; AB006550; AAK33945.1; -.  
 KW Transferase; Complete proteome.  
 SO SEQUENCE 163 AA; 17934 MW; B1ABB69B4C263262 CRC64;

Query Match 20.0%; Score 6; DB 16; Length 163;  
 Best Local Similarity 100.0%; Pred. No. 76;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 SLPELD 8  
 |||||  
 Db 139 SLPELD 144

```

RESULT 25
006112 PRELIMINARY; PRT; 184 AA.
AC 006112;
DT 01-JUN-1997 (TREMBlrel. 04, Created)
DT 01-JUN-1997 (TREMBlrel. 04, Last sequence update)
DT 01-AUG-1998 (TREMBlrel. 07, Last annotation update)
DE Hypothetical 21.5 kDa protein.
OS Methanococcus maripaludis.
OC Plasmid PUR8500.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanococcaceae; Methanococcus.
OX NCBI_TaxID=39152;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CS;
RX MEDLINE=97284503; PubMed=9139917;
RA Tumbula D.L., Bowen T.L., Whitman W.B.;
RT "Characterization of PUR8500 from the archaeon Methanococcus
maripaludis and construction of a shuttle vector.";
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CS;
RA Tumbula D.L., Whitman W.B.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U47023; AAC45262.1; -.
KM Hypothetical protein; Plasmid.
SQ SEQUENCE 184 AA; 21492 MW; 481963CF11BCC08 CRC64;

Query Match 20.0%; Score 6; DB 1; Length 184;
Best Local Similarity 100.0%; Pred. No. 84;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YSLPEL 7
DB 108 YSLPEL 113

RESULT 26
059673 PRELIMINARY; PRT; 189 AA.
AC 059673;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Superoxide dismutase (EC 1.15.1.1) (Fragment).
OS SOD.
OS Propionibacterium freudenreichii.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Propionibacteriaceae; Propionibacteriaceae;
OC Propionibacterium.
OX NCBI_TaxID=1744;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=P23;
RX MEDLINE=96074560; PubMed=7488202;
RA Gabblanelli R., Battistoni A., Polizio F., Carri M.T., De Martino A.,
Meier B., Desideri A., Rotilio G.;
RT "Metal uptake of recombinant cambialistic superoxide dismutase from
Propionibacterium shermanii is affected by growth conditions of host
Escherichia coli cells.";
RN Blochem. Biophys. Res. Commun. 216:841-847(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=P23;
RA Gabblanelli R.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE
CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS.
CC -1- CATALYTIC ACTIVITY: 2 PEROXIDE RADICAL + 2 H(+) = O(2) + H(2)O(2).
CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE
CC FAMILY.

```

```

DR EMBL; X91650; CAAG2838.1; -.
DR EMBL; Y09012; CAAT0215.1; -.
DR HSSP; P80293; IAWM.
DR InterPro; IPR001189; SODismutase.
DR Pfam; PF00081; sodef; 1.
DR Pfam; PF02777; sodef; C1.
DR ProDom; PD000475; SODismutase; 1.
DR PROSITE; PS00088; SOD_MN; 1.
KM Oxidoreductase.
FT NON_TER 189
SQ SEQUENCE 189 AA; 21334 MW; 1971419C716FF651 CRC64;

Query Match 20.0%; Score 6; DB 2; Length 189;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 EPTSG 20
DB 17 EPTSG 22

RESULT 27
090LUB PRELIMINARY; PRT; 193 AA.
ID 090LUB;
AC 090LUB;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Envelope glycoprotein (Fragment).
OS ENV.
OS Human immunodeficiency virus type 1.
OS Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1016;
RA Delwart E.L.;
RT "Homogeneous quasispecies in 16/17 individuals during very early HIV-1
primary infection.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF403533; AAK95241.1; -.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
KM AIDS; Coat protein; Glycoprotein.
FT NON_TER 1
FT NON_TER 193
SQ SEQUENCE 193 AA; 21366 MW; 1CD5BAA0920228B5 CRC64;

Query Match 20.0%; Score 6; DB 15; Length 193;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 ISGOIN 23
DB 155 ISGOIN 160

RESULT 28
09APY3 PRELIMINARY; PRT; 200 AA.
ID 09APY3;
AC 09APY3;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Manganese superoxide dismutase (EC 1.15.1.1).
OS SODA.
OS Corynebacterium melassecola, and
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacteriaceae; Corynebacteriaceae;
OC Corynebacterium.
OX NCBI_TaxID=41643, 1718;
RN [1]

```

RP SEQUENCE FROM N.A.  
 RC SPECIES=C.melasecola; STRAIN=ATCC17965;  
 RX MEDLINE=21101811; PubMed=11157941;  
 RA Merkum M., Guyonvarch A.;  
 RT "Cloning of the SODA Gene from *Corynebacterium melasecola* and Role of  
 RT Superoxide Dismutase in Cellular Viability."; *J. Bacteriol.* 183:1284-1295(2001).  
 RL [2]  
 RN SEQUENCE FROM N.A.  
 RP SPECIES=C.glutamicum; STRAIN=ATCC 13032;  
 RA Nakagawa S., Mizoguchi H., Ando S., Hayashi M., Hattori M., Shiba T.,  
 RA Sakaki Y., Yokoi H., Ozaki A.;  
 RT "SOD of *Corynebacterium glutamicum* ATCC 13032.";  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE  
 CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: 2 SUPEROXIDE + 2 H(+) = O(2) + H(2)O(2).  
 CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE  
 CC FAMILY.  
 CC EMBL; AF236111; AK01490.1; -  
 CC EMBL; AB055218; BAB62412.1; -  
 CC HSSP; P17670; IIDS  
 DR InterPro; IPR001189; SODismutase.  
 DR Pfam; PF00081; sode; 1.  
 DR ProDom; PD00475; SODismutase; 1.  
 DR PROSITE; PS00088; SOD\_MN; 1.  
 KW Oxidoreductase.  
 SQ SEQUENCE 200 AA; 22101 MW; D236A2C3F291CE13 CRC64;  
 Query Match 20.0%; Score 6; DB 2; Length 200;  
 Best Local Similarity 100.0%; Pred. No. 90;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 LPELDY 9  
 Db 6 LPELDY 11  
 RESULT 29  
 ID Q9YV92 PRELIMINARY; PRT; 200 AA.  
 AC Q9YV92;  
 DT 01-MAY-1999 (TREMBlrel. 10, Created)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE Envelope glycoprotein (Fragment).  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OC NCBI\_Taxid=11676;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=08-121V4;  
 RX MEDLINE=98248614; PubMed=9582120;  
 RA Frenkel L.M., Mullins J.I., Learn G.H., Manns-Arcuino L.,  
 RA Herring B.L., Kalish M.L., Steksee R.W., Thea D.M., Nichols J.E.,  
 RA Liu S.-L., Harmache M., He X., Muthui D., Madan A., Hood L.,  
 RA Haase A.T., Zupanic M., Steakus K., Molinsky S.M., Krogstad P.,  
 RA Zhao J.-Q., Chen I., Koup R., Ho D.D., Korber B.T., Apple R.J.;  
 RA Coombs R.W., Pahwa S., Roberts N.J. Jr.;  
 RT "Genetic evaluation of suspected cases of transient HIV-1 infection of  
 RT infants.";  
 RL Science 280:1073-1077(1998).  
 DR EMBL; AF065540; AAC21503.1; -  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 KW AIDS; Coat protein; Glycoprotein.  
 FT NON\_TER 1 200  
 FT SEQUENCE 200 AA; 22373 MW; 1844CAFA07BDBA8F6 CRC64;  
 Query Match 20.0%; Score 6; DB 15; Length 200;

Best Local Similarity 100.0%; Pred. No. 90;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 18 ISGOIN 23  
 Db 168 ISGOIN 173  
 RESULT 30  
 ID Q90009 PRELIMINARY; PRT; 200 AA.  
 AC Q90009;  
 DT 01-NOV-1998 (TREMBlrel. 08, Created)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE Envelope glycoprotein (Fragment).  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OC NCBI\_Taxid=11676;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=08-121V4;  
 RX MEDLINE=98248614; PubMed=9582120;  
 RA Frenkel L.M., Mullins J.I., Learn G.H., Manns-Arcuino L.,  
 RA Herring B.L., Kalish M.L., Steksee R.W., Thea D.M., Nichols J.E.,  
 RA Liu S.-L., Harmache M., He X., Muthui D., Madan A., Hood L.,  
 RA Haase A.T., Zupanic M., Steakus K., Molinsky S.M., Krogstad P.,  
 RA Zhao J.-Q., Chen I., Koup R., Ho D.D., Korber B.T., Apple R.J.,  
 RA Coombs R.W., Pahwa S., Roberts N.J. Jr.;  
 RT "Genetic evaluation of suspected cases of transient HIV-1 infection of  
 RT infants.";  
 RL Science 280:1073-1077(1998).  
 DR EMBL; AF065542; AAC21505.1; -  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 KW AIDS; Coat protein; Glycoprotein.  
 FT NON\_TER 1 200  
 FT SEQUENCE 200 AA; 22363 MW; 008F570236B1AC77 CRC64;  
 Query Match 20.0%; Score 6; DB 15; Length 200;  
 Best Local Similarity 100.0%; Pred. No. 90;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 18 ISGOIN 23  
 Db 168 ISGOIN 173  
 RESULT 31  
 ID Q70722 PRELIMINARY; PRT; 202 AA.  
 AC Q70722;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE Surface envelope glycoprotein (Fragment).  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OC NCBI\_Taxid=11676;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=91NE537;  
 RX MEDLINE=94365971; PubMed=8084001;  
 RA Delwart E.L., Sheppard H.W., Walker B.D., Goudsmut J., Mullins J.I.;  
 RT "Human immunodeficiency virus type 1 evolution in vivo tracked by DNA  
 RT heteroduplex mobility assays.";  
 RL J. Virol. 68:6672-6683(1994).  
 DR EMBL; U13243; AAA61888.1; -  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.

KM AIDS; Coat protein; Envelope protein; Glycoprotein.  
 FT NON\_TER 1  
 FT NON\_TER 202  
 SQ SEQUENCE 202 AA; 22474 MW; 9D2A804C9B93D75 CRC64;  
 Query Match  
 Best Local Similarity 100.0%; Score 6; DB 15; Length 202;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 18 ISGQIN 23  
 Db 167 ISGQIN 172

RESULT 32  
 ID Q70723 PRELIMINARY; PRT; 202 AA.  
 AC Q70723;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 RT "Human immunodeficiency virus type 1 evolution in vivo tracked by DNA  
 heteroduplex mobility assays."  
 RL J. Virol. 68:6672-6683(1994).  
 DR EMBL: U13244; AAA61889.1; -  
 DR InterPro: IPR000777; GP120.  
 DR Pfam: PF00516; GP120.1.  
 DE Envelope glycoprotein; Envelope protein; Glycoprotein.  
 FT NON\_TER 1  
 FT NON\_TER 202  
 SQ SEQUENCE 202 AA; 22399 MW; 83CB929A3F6A88F CRC64;  
 Query Match  
 Best Local Similarity 100.0%; Score 6; DB 15; Length 202;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 18 ISGQIN 23  
 Db 167 ISGQIN 172

RESULT 33  
 ID Q71823 PRELIMINARY; PRT; 202 AA.  
 AC Q71823;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 RT "Human immunodeficiency virus type 1."  
 RL J. Virol. 68:6672-6683(1994).  
 DR EMBL: U13244; AAA61889.1; -  
 DR InterPro: IPR000777; GP120.  
 DR Pfam: PF00516; GP120.1.  
 DE Envelope glycoprotein; Envelope protein; Glycoprotein.  
 FT NON\_TER 1  
 FT NON\_TER 202  
 SQ SEQUENCE 202 AA; 22399 MW; 83CB929A3F6A88F CRC64;  
 Query Match  
 Best Local Similarity 100.0%; Score 6; DB 15; Length 202;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 18 ISGQIN 23  
 Db 167 ISGQIN 172

RP SEQUENCE FROM N.A.  
 RC STRAIN=537 SAMPLE 1991;  
 RX MEDLINE=96089215; PubMed=8554905;  
 RA Delwart E.L., Busch M.P., Kalish M.L., Mosley J.W., Mullins J.I.;  
 RT "Rapid molecular epidemiology of human immunodeficiency virus  
 transmission."  
 RL AIDS Res. Hum. Retroviruses 11:1081-1093(1995).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=537 SAMPLE 1991;  
 RA Shpaer E.G.;  
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U23653; AAB0636.1; -  
 DR InterPro: IPR000777; GP120.  
 DR Pfam: PF00516; GP120.1.  
 DE AIDS; Coat protein; Glycoprotein.  
 FT NON\_TER 1  
 FT NON\_TER 202  
 SQ SEQUENCE 202 AA; 22474 MW; 9D2A804C9B93D75 CRC64;  
 Query Match  
 Best Local Similarity 100.0%; Score 6; DB 15; Length 202;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 18 ISGQIN 23  
 Db 167 ISGQIN 172

RESULT 34  
 ID Q71824 PRELIMINARY; PRT; 202 AA.  
 AC Q71824;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 RT "Human immunodeficiency virus type 1 evolution in vivo tracked by DNA  
 heteroduplex mobility assays."  
 RL J. Virol. 68:6672-6683(1994).  
 DR EMBL: U13244; AAA61889.1; -  
 DR InterPro: IPR000777; GP120.  
 DR Pfam: PF00516; GP120.1.  
 DE Envelope glycoprotein; Envelope protein; Glycoprotein.  
 FT NON\_TER 1  
 FT NON\_TER 202  
 SQ SEQUENCE 202 AA; 22399 MW; 83CB929A3F6A88F CRC64;  
 Query Match  
 Best Local Similarity 100.0%; Score 6; DB 15; Length 202;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 18 ISGQIN 23  
 Db 167 ISGQIN 172

RESULT 35  
 ID Q71824 PRELIMINARY; PRT; 202 AA.  
 AC Q71824;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 RT "Human immunodeficiency virus type 1 evolution in vivo tracked by DNA  
 heteroduplex mobility assays."  
 RL J. Virol. 68:6672-6683(1994).  
 DR EMBL: U13244; AAA61889.1; -  
 DR InterPro: IPR000777; GP120.  
 DR Pfam: PF00516; GP120.1.  
 DE Envelope glycoprotein; Envelope protein; Glycoprotein.  
 FT NON\_TER 1  
 FT NON\_TER 202  
 SQ SEQUENCE 202 AA; 22399 MW; 83CB929A3F6A88F CRC64;  
 Query Match  
 Best Local Similarity 100.0%; Score 6; DB 15; Length 202;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 18 ISGQIN 23  
 Db 167 ISGQIN 172

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 18 ISGQIN 23  
 |||||  
 Db 167 ISGQIN 172

## RESULT 35

ID 071827 PRELIMINARY; PRT; 202 AA.

AC 071827;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE Envelope glycoprotein, V3-V5 region (Fragment).  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
 OX NCBI\_Taxid=11676;

## SEQUENCE FROM N.A.

STRAIN=537 SAMPLE 1991;  
 MEDLINE=95194734; PubMed=7888227;  
 RA Shpaer E.G., Delwart E.L., Kulkarni C.L., Goudsmit J., Bachmann M.H.,  
 RA Mullins J.I.;  
 RT "Conserved V3 loop sequences and transmission of human  
 RT immunodeficiency virus type 1."  
 RL AIDS Res. Hum. Retroviruses 10:1679-1684(1994).  
 RN [2]

## SEQUENCE FROM N.A.

STRAIN=537 SAMPLE 1991;  
 RC MEDLINE=96089215; PubMed=8554905;  
 RA Delwart E.L., Buech M.P., Kallish M.L., Mosley J.W., Mullins J.I.;  
 RT "Rapid molecular epidemiology of human immunodeficiency virus  
 RT transmission."  
 RL AIDS Res. Hum. Retroviruses 11:1081-1093(1995).  
 RN [3]

## SEQUENCE FROM N.A.

STRAIN=537 SAMPLE 1991;  
 RC Shpaer E.G.;  
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U23657; AB0640.1; -.  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 KW AIDS; Coat protein; Glycoprotein.

## NON\_TER

FT NON\_TER 1  
 SQ SEQUENCE 202 AA; 22526 MW; 76AAA09215F29A39 CRC64;

## Query Match

Best Local Similarity 100.0%; Score 6; DB 15; Length 202;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 18 ISGQIN 23  
 |||||  
 Db 168 ISGQIN 173

## RESULT 36

ID 090010 PRELIMINARY; PRT; 203 AA.

AC 090010;  
 DT 01-NOV-1998 (TREMBlrel. 08, Created)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE Envelope glycoprotein (Fragment).  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
 OX NCBI\_Taxid=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=08-121V4;

FX MEDLINE=98248614; PubMed=9582120;  
 RA Frenkel L.M., Mullins J.I., Learn G.H., Manns-Arcuino L.,  
 RA Herring B.L., Kallish M.L., Steketee R.W., Thea D.M., Nichols J.E.,  
 RA Liu S.-L., Harmache A., He X., Muthi D., Madan A., Hood L.,  
 RA Haase A.T., Zupanic M., Steakus K., Wolinsky S.M., Krogstad P.,  
 RA Zhao J.-Q., Chen I., Koup R., Ho D.D., Korder B.T., Apple R.J.,  
 RA Coombs R.W., Pahwa S., Roberts N.J. Jr.;  
 RT "Genetic evaluation of suspected cases of transient HIV-1 infection of  
 RT infants."  
 RL Science 280:1073-1077(1998).  
 DR EMBL; AF065543; AAC21506.1; -.  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 KW AIDS; Coat protein; Glycoprotein.  
 FT NON\_TER 1  
 SQ SEQUENCE 203 AA; 22676 MW; 6C269D4172E9B74B CRC64;

## Query Match

Best Local Similarity 100.0%; Score 6; DB 15; Length 203;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 18 ISGQIN 23  
 |||||  
 Db 168 ISGQIN 173

## RESULT 37

ID 09YV91 PRELIMINARY; PRT; 203 AA.

AC 09YV91;  
 DT 01-MAY-1999 (TREMBlrel. 10, Created)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE Envelope glycoprotein (Fragment).  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
 OX NCBI\_Taxid=11676;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=08-121V4;  
 RX MEDLINE=98248614; PubMed=9582120;  
 RA Frenkel L.M., Mullins J.I., Learn G.H., Manns-Arcuino L.,  
 RA Herring B.L., Kallish M.L., Steketee R.W., Thea D.M., Nichols J.E.,  
 RA Liu S.-L., Harmache A., He X., Muthi D., Madan A., Hood L.,  
 RA Haase A.T., Zupanic M., Steakus K., Wolinsky S.M., Krogstad P.,  
 RA Zhao J.-Q., Chen I., Koup R., Ho D.D., Korder B.T., Apple R.J.,  
 RA Coombs R.W., Pahwa S., Roberts N.J. Jr.;  
 RT "Genetic evaluation of suspected cases of transient HIV-1 infection of  
 RT infants."  
 RL Science 280:1073-1077(1998).  
 DR EMBL; AF065544; AAC21507.1; -.  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 KW AIDS; Coat protein; Glycoprotein.  
 FT NON\_TER 1  
 SQ SEQUENCE 203 AA; 22602 MW; 8C6C8495742F6E35 CRC64;

## Query Match

Best Local Similarity 100.0%; Score 6; DB 15; Length 203;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 18 ISGQIN 23  
 |||||  
 Db 168 ISGQIN 173

## RESULT 38

ID 09YV90 PRELIMINARY; PRT; 203 AA.

AC 09YV90;  
 DT 01-MAY-1999 (TREMBlrel. 10, Created)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE Envelope glycoprotein (Fragment).  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
 OX NCBI\_Taxid=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=08-121V4;

## Query Match

Best Local Similarity 100.0%; Score 6; DB 15; Length 203;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 18 ISGQIN 23  
 |||||  
 Db 168 ISGQIN 173



```
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OC NCBI_Taxid=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=08-121V4;
RX MEDLINE=98248614; Pubmed=9582120;
RA Frenkel L.M., Mullins J.I., Learn G.H., Manns-Arcuino L.,
RA Herring B.L., Kalish M.L., Steketee R.W., Thea D.M., Nichols J.E.,
RA Liu S.-L., Harmache A., He X., Muchi D., Madan A., Hood L.,
RA Haase A.T., Zupnick M., Staskus K., Molinsky S.M., Krosgard P.,
RA Zhao J.-O., Chen I., Koup R., Ho D.D., Korber B.T., Apple R.J.,
RA Combs R.W., Pahwa S., Roberts N.J. Jr.;
RT "Genetic evaluation of suspected cases of transient HIV-1 infection of
RT infants.";
DR EMBL: AF065546; AAC1509.1; -.
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
KW AIDS; Coat protein; Glycoprotein.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 203 AA; 22562 MW; 6130090845823D4 CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 15; Length 203;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 ISGOIN 23
Db 168 ISGOIN 173

RESULT 39
ID Q70724 PRELIMINARY; PRT; 203 AA.
AC Q70724;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Surface envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OC NCBI_Taxid=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=91NE537;
RX MEDLINE=94365971; Pubmed=8084001;
RA Delwart E.L., Sheppard H.W., Walker B.D., Goudsmit J., Mullins J.I.;
RT "Human immunodeficiency virus type 1 evolution in vivo tracked by DNA
RT heteroduplex mobility assays.";
RT J. Virol. 68:6672-6683(1994).
DR EMBL: U13245; AAA61890.1; -.
DR InterPro: IPR00777; GP120.
DR Pfam: PF00516; GP120; 1.
KW AIDS; Coat protein; Envelope protein; Glycoprotein.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 203 AA; 22637 MW; CA202D16E061AB96 CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 15; Length 203;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 ISGOIN 23
Db 168 ISGOIN 173
```

```
RESULT 40
ID Q70725 PRELIMINARY; PRT; 203 AA.
AC Q70725;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Surface envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OC NCBI_Taxid=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=91NE537;
RX MEDLINE=94365971; Pubmed=8084001;
RA Delwart E.L., Sheppard H.W., Walker B.D., Goudsmit J., Mullins J.I.;
RT "Human immunodeficiency virus type 1 evolution in vivo tracked by DNA
RT heteroduplex mobility assays.";
RT J. Virol. 68:6672-6683(1994).
DR EMBL: U13246; AAA61891.1; -.
DR InterPro: IPR00777; GP120.
DR Pfam: PF00516; GP120; 1.
KW AIDS; Coat protein; Envelope protein; Glycoprotein.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 203 AA; 22686 MW; 4C325CB1DA45BADE CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 15; Length 203;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 ISGOIN 23
Db 168 ISGOIN 173

Search completed: April 9, 2003, 14:16:49
Job time : 40.8624 secs
```



GenCore version 5.1.4 p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 9, 2003, 14:11:20 ; Search time 7.29358 Seconds  
(without alignments)  
395.421 Million cell updates/sec

Title: US-09-987-190-2

Perfect score: 30

Sequence: 1 KYSLEPDLDFEFSATEPYSQINEIXYTX 30

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283224 seqs, 96134422 residues

Database size : 0

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10	33.3	233	1 DSBVN	superoxide dismuta
2	7	23.3	207	2 JC4351	superoxide dismuta
3	7	23.3	207	2 S15205	superoxide dismuta
4	7	23.3	207	2 S60669	superoxide dismuta
5	7	23.3	328	2 T03296	beta-glucosidase (
6	7	23.3	488	2 B69959	glycine dehydrogen
7	7	23.3	507	2 T10791	beta-glucosidase (
8	7	23.3	531	2 S23940	beta-glucosidase (
9	7	23.3	1034	2 S76134	hypothetical prote
10	7	23.3	2055	2 T31617	hypothetical prote
11	6	20.0	91	2 AH1065	hypothetical prote
12	6	20.0	202	2 JC4396	superoxide dismuta
13	6	20.0	207	2 B86498	superoxide dismuta
14	6	20.0	207	2 B72124	superoxide dismuta
15	6	20.0	207	2 AH1190	3-methyladenine DN
16	6	20.0	278	2 C83587	hypothetical prote
17	6	20.0	278	2 E64131	formyltetrahydrofo
18	6	20.0	293	2 T06212	glucose and ribito
19	6	20.0	334	2 H87192	probable isomerase
20	6	20.0	338	2 D86815	hypothetical prote
21	6	20.0	342	2 E71359	conserved hypotet
22	6	20.0	345	2 T25408	hypothetical prote
23	6	20.0	352	2 D84329	heme biosynthesis
24	6	20.0	355	2 G83003	A / G specific ade
25	6	20.0	379	2 T08277	carotenoid biosynt
26	6	20.0	384	2 G84650	probable beta-gluc
27	6	20.0	385	2 F70551	probable ketB prot
28	6	20.0	385	2 G87006	probable transmemb
29	6	20.0	405	2 A32812	repa protein - Agr

30	6	20.0	405	2 A13232	replication protei
31	6	20.0	412	2 D96939	sensory protein, c
32	6	20.0	428	2 C64080	hemy protein homol
33	6	20.0	429	2 T50362	cdc7-like protein
34	6	20.0	441	2 AH2930	oxidoreductase Atu
35	6	20.0	443	2 T31101	probable phosphom
36	6	20.0	446	2 F98351	probable oxidoredu
37	6	20.0	467	2 S30839	UTR2 protein - yea
38	6	20.0	482	2 E97012	probable non-proce
39	6	20.0	488	2 F84001	glycine dehydrogen
40	6	20.0	506	2 T02400	probable beta-gluc
41	6	20.0	508	2 AD0076	probable membrane
42	6	20.0	514	2 T47837	beta-glucosidase-1
43	6	20.0	517	2 T02403	probable beta-gluc
44	6	20.0	535	2 S40461	t-complex-type mol
45	6	20.0	535	2 S40462	t-complex-type mol
46	6	20.0	545	2 T47952	hypothetical prote
47	6	20.0	551	2 T09647	beta-glucosidase (
48	6	20.0	551	2 C64903	probable ABC-type
49	6	20.0	561	2 D85726	hypothetical prote
50	6	20.0	561	2 E90891	hypothetical prote
51	6	20.0	588	2 B95406	probable oxidoredu
52	6	20.0	617	2 T03510	hypothetical prote
53	6	20.0	640	2 A86657	fructose-bisphosph
54	6	20.0	647	2 G72346	hypothetical prote
55	6	20.0	661	2 D69907	hypothetical prote
56	6	20.0	699	2 T11578	probable lipoxigen
57	6	20.0	948	2 B81883	excinuclease ABC c
58	6	20.0	949	2 A81138	excinuclease ABC c
59	6	20.0	965	2 C82560	excinuclease ABC c
60	6	20.0	1146	2 S59376	hypothetical prote
61	6	20.0	1679	2 T50091	Yeast Ecm29 cell w
62	6	20.0	1687	2 T30244	phosphodiesterase
63	6	20.0	1706	2 T30175	exoribonuclease, v
64	6	20.0	1719	2 T30174	exoribonuclease, v
65	6	20.0	1822	2 S63985	collagen alpha 2 c
66	6	20.0	3198	2 A43426	collagen alpha 2 f
67	5	16.7	6	2 I51317	BMH transcription
68	5	16.7	38	2 C60657	circumsporozoit p
69	5	16.7	47	2 AF0727	conserved hypotet
70	5	16.7	49	2 S50999	superoxide dismuta
71	5	16.7	54	2 AB1819	hypothetical prote
72	5	16.7	70	2 F81899	hypothetical prote
73	5	16.7	76	2 T29961	hypothetical prote
74	5	16.7	85	2 S44604	C02F5.2 protein -
75	5	16.7	85	2 AB1868	hypothetical prote
76	5	16.7	88	2 AH2013	hypothetical prote
77	5	16.7	91	2 I54248	apolipoprotein A-I
78	5	16.7	101	2 S39490	hypothetical prote
79	5	16.7	101	2 H72556	hypothetical prote
80	5	16.7	103	2 AD1172	conserved hypotet
81	5	16.7	107	2 H72506	hypothetical prote
82	5	16.7	109	2 C82913	conserved hypotet
83	5	16.7	115	2 S15822	naturiletic peptid
84	5	16.7	115	2 T38514	small nuclear ribo
85	5	16.7	119	2 G85846	unknown protein en
86	5	16.7	119	2 AD2513	hypothetical prote
87	5	16.7	120	2 T07888	protein kinase (BC
88	5	16.7	122	2 S75687	ferredoxin (2Fe-2S
89	5	16.7	122	2 B86316	protein T10022.5 l
90	5	16.7	124	1 VKLJSI	trans-regulatory s
91	5	16.7	124	2 F75373	conserved hypotet
92	5	16.7	125	2 E70332	conserved hypotet
93	5	16.7	125	2 C75305	conserved hypotet
94	5	16.7	127	2 T10311	hypothetical prote
95	5	16.7	128	2 B70910	hypothetical prote
96	5	16.7	129	1 QOCWV4	AC2 protein - abut
97	5	16.7	129	1 JQ1871	AC2 protein - toma
98	5	16.7	129	2 T19154	hypothetical prote
99	5	16.7	132	2 TQ2302	Al3 protein - pepp
100	5	16.7	132	2 S31877	Al3 protein - pepp
101	5	16.7	133	2 H82366	conserved hypotet
102	5	16.7	134	2 A12102	hypothetical prote

replication protei  
sensory protein, c  
hemy protein homol  
cdc7-like protein  
oxidoreductase Atu  
probable phosphom  
probable oxidoredu  
UTR2 protein - yea  
probable non-proce  
glycine dehydrogen  
probable beta-gluc  
probable membrane  
beta-glucosidase-1  
probable beta-gluc  
t-complex-type mol  
t-complex-type mol  
hypothetical prote  
beta-glucosidase (

103	5	16.7	135	2	B87516	conserved hypotet
104	5	16.7	135	2	AH1093	hypothetical prote
105	5	16.7	137	2	S74181	tumor antigen UK11
106	5	16.7	137	2	S39853	trkB protein - Str
107	5	16.7	138	2	D82400	TonB system transp
108	5	16.7	141	2	G91001	hypothetical prote
109	5	16.7	143	2	D90110	40S ribosomal prot
110	5	16.7	145	2	S41221	hypothetical prote
111	5	16.7	147	2	A75093	hypothetical prote
112	5	16.7	147	2	F71050	hypothetical prote
113	5	16.7	148	2	T33812	hypothetical prote
114	5	16.7	149	2	US0275	arginine metabolis
115	5	16.7	149	2	A83997	transcription regu
116	5	16.7	149	2	AC1608	arginine repressor
117	5	16.7	149	2	AG1245	hypothetical prote
118	5	16.7	150	2	F81280	hypothetical prote
119	5	16.7	152	2	S76383	hypothetical prote
120	5	16.7	154	2	CS3373	hypothetical prote
121	5	16.7	154	2	D83768	bacteriophage phi C
122	5	16.7	157	2	E84997	16 kD heat shock p
123	5	16.7	159	2	A70833	carbon-monoxide de
124	5	16.7	161	2	S49488	collagen alpha 3(I
125	5	16.7	162	2	S28301	hypothetical prote
126	5	16.7	162	2	F86637	hypothetical prote
127	5	16.7	164	2	T50870	hypothetical prote
128	5	16.7	165	2	PS0261	calcineurin regula
129	5	16.7	165	2	I40515	spal protein - Bac
130	5	16.7	165	2	S75265	hypothetical prote
131	5	16.7	165	2	A97759	16S rRNA processin
132	5	16.7	166	2	T13192	hypothetical prote
133	5	16.7	168	2	B75498	conserved hypotet
134	5	16.7	168	2	B86435	protein F17F8.17 f
135	5	16.7	169	2	CS9394	hypothetical prote
136	5	16.7	170	1	A33391	calcineurin regula
137	5	16.7	170	1	S34127	calcineurin regula
138	5	16.7	170	1	UC1220	calcineurin regula
139	5	16.7	170	2	AA4307	calcineurin regula
140	5	16.7	170	2	JCS174	calcineurin regula
141	5	16.7	170	2	AC2917	invasion protein A
142	5	16.7	171	2	B69116	conserved hypotet
143	5	16.7	171	2	AH3585	hypothetical prote
144	5	16.7	175	2	T11151	NADH2 dehydrogenas
145	5	16.7	175	2	T11413	NADH2 dehydrogenas
146	5	16.7	175	2	D75467	bacteriophage phi C
147	5	16.7	176	1	S02511	flavodoxin - Klebs
148	5	16.7	176	2	J01232	calcineurin regula
149	5	16.7	177	2	B55517	hypothetical prote
150	5	16.7	178	2	F90168	hypothetical prote
151	5	16.7	179	2	C97711	hypothetical prote
152	5	16.7	180	2	S38955	alpha-amylose inh1
153	5	16.7	180	2	F71639	hypothetical prote
154	5	16.7	181	2	T49476	hypothetical prote
155	5	16.7	183	2	B69388	hypothetical prote
156	5	16.7	184	2	B82307	16S rRNA processin
157	5	16.7	185	1	AA4559	relaxin precursor
158	5	16.7	185	2	S48082	relaxin precursor
159	5	16.7	185	2	I39736	hypothetical prote
160	5	16.7	185	2	AE1902	hypothetical prote
161	5	16.7	188	2	T16749	hypothetical prote
162	5	16.7	189	2	T40699	hypothetical prote
163	5	16.7	190	2	A26630	hypothetical prote
164	5	16.7	191	2	T09985	protoporphyrinogen
165	5	16.7	191	2	F97691	hypothetical prote
166	5	16.7	195	2	H70171	hypothetical prote
167	5	16.7	195	2	G77152	hypothetical prote
168	5	16.7	196	2	G89950	conserved hypotet
169	5	16.7	196	2	AD2304	hypothetical prote
170	5	16.7	199	2	J01398	hypothetical prote
171	5	16.7	200	2	A26169	nucleoplasmic - Af
172	5	16.7	200	2	T02173	probable alcohol d
173	5	16.7	202	2	S51097	superoxide dismuta
174	5	16.7	203	2	S04860	alpha-amylose/subt
175	5	16.7	203	2	AB1912	peptidyl-tyrRNA hydr
176	5	16.7	204	2	S22639	alpha-amylose/subt
177	5	16.7	205	2	E82601	conserved hypotet
178	5	16.7	206	2	H84165	hypothetical prote
179	5	16.7	207	2	G69835	hypothetical prote
180	5	16.7	208	2	H64139	guanylate kinase (
181	5	16.7	208	2	H69522	conserved hypotet
182	5	16.7	209	2	C96999	hypothetical prote
183	5	16.7	210	2	H82037	conserved hypotet
184	5	16.7	210	2	A99487	hypothetical prote
185	5	16.7	211	2	E90174	superoxide dismuta
186	5	16.7	211	2	S34616	superoxide dismuta
187	5	16.7	211	2	T42603	gene 60 protein -
188	5	16.7	211	2	S15167	fixt protein - Azo
189	5	16.7	211	2	H83239	pseudouridine synt
190	5	16.7	212	2	WZBEP2	gene 60 protein -
191	5	16.7	212	2	F84862	probable extensin
192	5	16.7	214	2	AB2157	hypothetical prote
193	5	16.7	215	2	S39542	probable glutathio
194	5	16.7	215	2	F72641	hypothetical prote
195	5	16.7	216	1	S42716	calcineurin regula
196	5	16.7	216	2	S37349	luxM protein - Vib
197	5	16.7	218	2	H83485	conserved hypotet
198	5	16.7	218	2	AG0883	conserved hypotet
199	5	16.7	219	2	C81038	phosphoglycolate p
200	5	16.7	220	2	B49736	collagen alpha 3(I
201	5	16.7	220	2	H70568	hypothetical prote
202	5	16.7	221	2	T13935	envelope protein -
203	5	16.7	223	2	F83034	urease accessory p
204	5	16.7	224	2	A53143	testis-determining
205	5	16.7	225	2	H75088	phenomone shuttlin
206	5	16.7	225	2	T20250	hypothetical prote
207	5	16.7	226	1	S19715	nitrite hydratase
208	5	16.7	226	2	A35968	tox regulon transc
209	5	16.7	226	2	AI2774	ribose 5-phosphate
210	5	16.7	226	2	AF2362	transcription regu
211	5	16.7	228	2	H97336	uncharacterized co
212	5	16.7	230	2	B86824	two-component syst
213	5	16.7	230	2	S76832	hypothetical prote
214	5	16.7	231	2	G97554	ribose 5-phosphate
215	5	16.7	231	2	H64850	flagellar hook for
216	5	16.7	231	2	E90810	flagellar biosynth
217	5	16.7	231	2	A85670	hypothetical prote
218	5	16.7	231	2	G75451	conserved hypotet
219	5	16.7	232	2	S47641	flagellar hook for
220	5	16.7	232	2	AI0639	flagellar hook for
221	5	16.7	234	1	S76823	hypothetical prote
222	5	16.7	236	2	F64964	yeep protein - Bsc
223	5	16.7	236	2	B84812	hypothetical prote
224	5	16.7	237	2	D97852	hypothetical prote
225	5	16.7	237	2	T22859	hypothetical prote
226	5	16.7	237	2	AI2062	hypothetical prote
227	5	16.7	238	2	AC2043	heme oxygenase (im
228	5	16.7	239	2	G82834	D-ribulose-5-phosp
229	5	16.7	239	2	F82386	amino acid ABC tra
230	5	16.7	239	2	C65222	probable 2-compone
231	5	16.7	239	2	B91267	two-component tran
232	5	16.7	239	2	G86107	probable 2-compone
233	5	16.7	239	2	F96908	uncharacterized co
234	5	16.7	241	2	T25506	hypothetical prote
235	5	16.7	241	2	T26676	hypothetical prote
236	5	16.7	242	2	T46543	hypothetical prote
237	5	16.7	242	2	S48568	DNA-directed DNA p
238	5	16.7	243	1	IOECO	conserved hypotet
239	5	16.7	243	1	E64452	DNA polymerase III
240	5	16.7	243	2	C85506	DNA polymerase III
241	5	16.7	243	2	C85506	DNA polymerase III
242	5	16.7	244	2	H70905	DNA polymerase III
243	5	16.7	244	2	S38568	transcription init
244	5	16.7	245	1	A44914	RNA nucleotidyltr
245	5	16.7	245	2	B70480	ribosomal protein
246	5	16.7	246	2	I48302	collagen alpha 3(I
247	5	16.7	247	2	AG2138	hypothetical prote
248	5	16.7	248	2	AG0966	conserved hypotet
249	5	16.7	248	2	A97486	prokaryotic type I

249	5	16.7	248	2	A12703	prokaryotic type I	322	5	16.7	303	2	B84744	hypothetical prote
250	5	16.7	248	2	E89820	hypothetical prote	323	5	16.7	303	2	T46405	hypothetical prote
251	5	16.7	248	2	F69198	conserved hypochet	324	5	16.7	304	2	E90521	hypothetical prote
252	5	16.7	250	2	A71268	probable tRNA (gua	325	5	16.7	305	1	IDECRK	replication initia
253	5	16.7	250	2	B87921	protein P56G4.5 [i	326	5	16.7	306	2	T06297	hypothetical prote
254	5	16.7	251	1	BVECBK	biotin biosynthesi	327	5	16.7	309	2	B72342	conserved hypochet
255	5	16.7	251	2	G90735	probable enzyme B1	328	5	16.7	310	2	B86825	aspartate carboxy
256	5	16.7	251	2	H85585	hypothetical prote	329	5	16.7	310	2	S07522	capsid assembly pr
257	5	16.7	251	2	AB0597	biotin synthesis	330	5	16.7	310	2	AC3238	hypothetical prote
258	5	16.7	252	2	AG3472	phosphoadenylyl-su	331	5	16.7	311	2	D69881	yabc protein homol
259	5	16.7	253	2	S48430	probable membrane	332	5	16.7	312	2	E90929	probable flavoprot
260	5	16.7	254	2	E71203	hypothetical prote	333	5	16.7	312	2	A85778	probable flavoprot
261	5	16.7	255	2	JG0179	superoxide dismuta	334	5	16.7	312	2	B64928	electron transfer
262	5	16.7	256	2	T51150	probable transmemb	335	5	16.7	312	2	T23646	hypothetical prote
263	5	16.7	257	1	QUECFP	transfer protein t	336	5	16.7	312	2	B64072	spermidine/putresc
264	5	16.7	257	2	AB1297	ABC transporter (A	337	5	16.7	312	2	H97342	ABC-type MDR trans
265	5	16.7	257	2	AH1668	ABC transporter (A	338	5	16.7	312	2	T50705	transcription fact
266	5	16.7	257	2	T40655	hypothetical prote	339	5	16.7	313	2	E90634	probable flavoprot
267	5	16.7	258	2	T02176	probable alcohol d	340	5	16.7	313	2	E85485	probable flavoprot
268	5	16.7	258	2	A10566	hydroxypyruvate is	341	5	16.7	313	2	AF0511	FlxB protein [lipo
269	5	16.7	259	2	D64362	hypothetical prote	342	5	16.7	313	2	B64725	electron transfer
270	5	16.7	260	2	D24584	spectinomycin resi	343	5	16.7	313	2	S60713	polygalacturonase-
271	5	16.7	260	2	AB2469	exodeoxyribonuclea	344	5	16.7	314	2	H83321	conserved hypochet
272	5	16.7	260	2	JQ2330	Brl protein - Ind1	345	5	16.7	314	2	C75003	na+/ca+ exchanging
273	5	16.7	260	2	T50725	reaction center pr	346	5	16.7	315	1	HQDVSB	cytochrome-c3 hydr
274	5	16.7	260	2	A26538	reaction center pr	347	5	16.7	315	2	T37901	probable purine nu
275	5	16.7	261	2	A34476	collagen alpha 2(I	348	5	16.7	315	2	A91118	probable ferrichro
276	5	16.7	262	2	AB2488	hypothetical prote	349	5	16.7	315	2	H85962	probable ferrichro
277	5	16.7	263	2	AD2517	hypothetical prote	350	5	16.7	316	2	C64963	transcription regu
278	5	16.7	267	2	AB0615	SmtA protein [lipo	351	5	16.7	316	2	E85823	transcription regu
279	5	16.7	268	2	D75574	hypothetical prote	352	5	16.7	316	2	G90976	transcription regu
280	5	16.7	269	2	I51539	MHC class II beta	353	5	16.7	316	2	T15896	hypothetical prote
281	5	16.7	270	1	AC1956	phosphoenolpyruvate	354	5	16.7	318	2	F81205	chiamine-phosphat
282	5	16.7	270	2	F86177	protein Flp19.1 [	355	5	16.7	318	2	H81781	hypothetical prote
283	5	16.7	271	2	AC0157	probable short cha	356	5	16.7	318	2	A83708	hypothetical prote
284	5	16.7	271	2	F82261	hypothetical prote	357	5	16.7	319	2	H69364	hypothetical prote
285	5	16.7	272	2	E64248	hypothetical prote	358	5	16.7	320	2	AG0934	bifunctional prote
286	5	16.7	273	2	AC0294	Conserved hypochet	359	5	16.7	320	2	G70365	glucose-6-phosphat
287	5	16.7	273	2	T12793	hypothetical prote	360	5	16.7	320	2	B72656	hypothetical prote
288	5	16.7	274	2	G83186	hypothetical prote	361	5	16.7	321	1	BVECBP	biotin-lacetyl-CoA
289	5	16.7	274	2	P81250	probable membrane	362	5	16.7	321	2	D91241	hypothetical prote
290	5	16.7	276	2	S56550	hypothetical 30.6K	363	5	16.7	321	2	A86089	hypothetical prote
291	5	16.7	276	2	D91289	hypothetical prote	364	5	16.7	322	2	D64996	putative glycosyl
292	5	16.7	276	2	G86130	hypothetical prote	365	5	16.7	322	2	F91021	probable sugar tra
293	5	16.7	276	2	D84547	NM (no apical mer	366	5	16.7	322	2	E85865	probable sugar tra
294	5	16.7	276	2	T46368	hypothetical prote	367	5	16.7	322	2	E70306	conserved hypochet
295	5	16.7	278	2	A11977	hypothetical prote	368	5	16.7	323	2	A96670	hypothetical prote
296	5	16.7	278	2	T31935	hypothetical prote	369	5	16.7	323	2	H89923	hypothetical prote
297	5	16.7	279	2	B70180	conserved hypochet	370	5	16.7	324	2	A10434	probable sodium/ca
298	5	16.7	280	2	AG0987	conserved hypochet	371	5	16.7	325	2	C91138	hypothetical prote
299	5	16.7	281	2	C69709	superoxide dismuta	372	5	16.7	325	2	H70773	probable protoporp
300	5	16.7	283	2	B98257	lysR-type transcri	373	5	16.7	325	2	F85983	hypothetical prote
301	5	16.7	283	2	AI3027	transcription regu	374	5	16.7	325	2	A10904	probable membrane
302	5	16.7	285	2	AG1045	hypothetical prote	375	5	16.7	326	2	F65110	hypothetical 34.7
303	5	16.7	286	2	H90066	hypothetical prote	376	5	16.7	326	2	H83187	hypothetical prote
304	5	16.7	287	2	G90301	hypothetical prote	377	5	16.7	326	2	S76400	hypothetical prote
305	5	16.7	288	2	A87051	probable protoporp	378	5	16.7	326	2	D70548	probable muconate
306	5	16.7	288	2	T12462	hypothetical prote	379	5	16.7	327	2	D10673	probable lipopolys
307	5	16.7	289	2	A12030	yopt translocation	380	5	16.7	327	2	A10793	polygalacturonase
308	5	16.7	289	2	E86593	yopt translocation	381	5	16.7	328	2	H90193	hypothetical prote
309	5	16.7	290	2	C90803	hypothetical prote	382	5	16.7	328	2	T41074	hypothetical prote
310	5	16.7	290	2	G85611	probable histone Z	383	5	16.7	332	2	T35917	probable regulator
311	5	16.7	293	2	AF0617	probable DNA methy	384	5	16.7	332	2	B98258	ribosome ABC transpo
312	5	16.7	295	2	AH1773	transcription regu	385	5	16.7	332	2	AG3026	hypothetical prote
313	5	16.7	297	2	E86461	F14M2.8 protein -	386	5	16.7	332	2	H82064	hypothetical prote
314	5	16.7	298	2	UC7323	cyclic ADP-ribose-	387	5	16.7	333	1	QQVZF9	conserved hypochet
315	5	16.7	298	2	H97072	zn-dependent hydro	388	5	16.7	333	2	F42513	poly(A) polymerase
316	5	16.7	299	2	T39602	conserved hypochet	389	5	16.7	333	2	E36845	L3R protein - vari
317	5	16.7	300	2	A70433	hypothetical prote	390	5	16.7	333	2	T28518	hypothetical prote
318	5	16.7	300	2	AE2088	hypothetical prote	391	5	16.7	333	2	F72160	MAR protein - vari
319	5	16.7	301	2	C86666	flu operon transcr	392	5	16.7	334	2	C95272	conserved hypochet
320	5	16.7	302	2	C64311	Na+/Ca2+-exchangin	393	5	16.7	337	2	AF0086	periplasmic bindin
321	5	16.7	303	2	A69749	hypothetical prote	394	5	16.7	337	2	T43372	Chromo domain prot

395	5	16.7	338	2	E95244	glycerol-3-phospha
396	5	16.7	338	2	C98109	glycerol-3-phospha
397	5	16.7	338	2	E71872	hypothetical prote
398	5	16.7	339	2	D83201	probable glycosyl
399	5	16.7	341	2	D82852	biotin synthase (E
400	5	16.7	343	2	AC2075	iron(III) dicitrat
401	5	16.7	343	2	C75260	conserved hypotnet
402	5	16.7	343	2	A57253	chaperone-like pro
403	5	16.7	343	2	T32232	hypothetical prote
404	5	16.7	344	1	E11776	hydrogenase (Ec 1.
405	5	16.7	344	2	E83445	hypothetical prote
406	5	16.7	346	2	H70860	alcohol dehydrogen
407	5	16.7	346	2	UC1376	alcohol dehydrogen
408	5	16.7	346	2	G83127	conserved hypotnet
409	5	16.7	347	2	T39550	probable quinone o
410	5	16.7	347	2	G86675	carotenoid biosynt
411	5	16.7	347	2	VQ2359	wheat aluminum ind
412	5	16.7	348	2	AB1496	polyol (sorbitol)
413	5	16.7	348	2	AC1138	polyol (sorbitol)
414	5	16.7	348	2	S23900	nitrogen regulatio
415	5	16.7	349	1	RGEBCG	nitrogen regulatio
416	5	16.7	349	1	QGBED5	HHRF5 protein - hu
417	5	16.7	349	2	SS3023	nitrogen regulatio
418	5	16.7	349	2	A24114	nitrogen regulatio
419	5	16.7	349	2	AB0950	Two-component syst
420	5	16.7	349	2	AH0003	Two-component regu
421	5	16.7	349	2	F86074	histidine protein
422	5	16.7	350	2	S76698	hypothetical prote
423	5	16.7	351	2	AC0758	Chlg protein [impo
424	5	16.7	351	2	AC0758	v-type ATP synthas
425	5	16.7	352	2	G75487	genome polyprotein
426	5	16.7	355	2	S47219	probable sugar upt
427	5	16.7	355	2	E95990	3-dehydroquinatase
428	5	16.7	356	2	C97010	SMK1 protein - Pod
429	5	16.7	356	2	S39889	hypothetical prote
430	5	16.7	357	2	H71122	hypothetical prote
431	5	16.7	357	2	C89880	zinc metalloendope
432	5	16.7	357	2	B75578	conserved hypotnet
433	5	16.7	357	2	C95960	hydrogenase (Ec 1.
434	5	16.7	359	1	JO0805	hypothetical prote
435	5	16.7	359	2	D69111	hydrogenase (Ec 1.
436	5	16.7	360	1	S11968	P3 protein - arabi
437	5	16.7	360	1	P3VVM	1lipopolysaccharide
438	5	16.7	360	2	B72359	conserved hypotnet
439	5	16.7	360	2	B87286	alcohol dehydrogen
440	5	16.7	362	2	D87125	3-dehydroquinatase
441	5	16.7	362	2	G64054	hydrogenase (Ec 1.
442	5	16.7	363	1	HQZJUS	cbid protein - Syn
443	5	16.7	364	1	S77360	iron (iii) abc tra
444	5	16.7	366	2	S53898	probable membrane
445	5	16.7	366	2	A75077	hypothetical prote
446	5	16.7	367	2	E69998	YopM protein - Yer
447	5	16.7	367	2	A33950	N-acetylmuramoyl-L
448	5	16.7	367	2	T12790	hypothetical prote
449	5	16.7	368	2	C72260	hypothetical prote
450	5	16.7	368	2	T22587	hypothetical prote
451	5	16.7	369	2	B71557	probable adenine g
452	5	16.7	369	2	T22708	hypothetical prote
453	5	16.7	370	2	D84846	hypothetical prote
454	5	16.7	370	2	C83120	probable RND efflu
455	5	16.7	371	2	S60903	hypothetical prote
456	5	16.7	371	2	T49908	hypothetical prote
457	5	16.7	372	1	G69553	3',5'-cyclic-nucle
458	5	16.7	372	2	S51137	ATP-dependent 26S
459	5	16.7	372	2	S77549	patA2 protein - Sy
460	5	16.7	372	2	A53050	breifeldin A estera
461	5	16.7	372	2	AB2061	glycoyltransferas
462	5	16.7	374	2	B88192	hypothetical prote
463	5	16.7	374	2	C86039	hypothetical prote
464	5	16.7	374	2	AC0972	lipopolysaccharide
465	5	16.7	374	2	B42595	glycoyltransferas
466	5	16.7	374	2	S38674	recombination prot
467	5	16.7	374	2	T35139	recA protein - Str
468	5	16.7	374	2	AH1903	hypothetical prote
469	5	16.7	375	2	S75182	glutamate 5-kinase
470	5	16.7	375	2	B40205	Na+/H+-exchanging
471	5	16.7	376	2	TS1170	homocitrate syntha
472	5	16.7	377	2	F96550	hypothetical prote
473	5	16.7	378	2	E82839	sugar transferase
474	5	16.7	379	2	C90987	probable polysacch
475	5	16.7	379	2	F85632	probable polysacch
476	5	16.7	379	2	AC0770	probable polysacch
477	5	16.7	379	2	E64972	probable polysacch
478	5	16.7	379	2	H64117	spermidine/putresc
479	5	16.7	380	2	T11381	ubiquinol-cytochro
480	5	16.7	381	2	A72601	hypothetical prote
481	5	16.7	381	2	AF0320	probable N-acetyl g
482	5	16.7	383	2	T08705	hypothetical prote
483	5	16.7	385	2	T18821	hypothetical prote
484	5	16.7	385	2	H97498	permease AGR C 211
485	5	16.7	385	2	AE2717	permease [imported
486	5	16.7	385	2	AH2200	Two-component hybr
487	5	16.7	388	2	TS0078	hypothetical prote
488	5	16.7	389	2	D82223	ribonucleoside-dip
489	5	16.7	389	2	A48966	methylase - Lactoc
490	5	16.7	390	2	B40362	tryptophan syntha
491	5	16.7	390	2	S52015	catechol O-methyl
492	5	16.7	393	2	A10319	probable 2-octapre
493	5	16.7	393	2	D75418	homocitrate syntha
494	5	16.7	393	2	S67763	probable membrane
495	5	16.7	394	2	A56115	dual specificity p
496	5	16.7	395	2	A86166	protein F21B7.6 [i
497	5	16.7	395	2	TO1392	leucine-rich repa
498	5	16.7	396	1	JOH631	leucine-rich repa
499	5	16.7	396	1	LPNUH4	apolipoprotein A-I
500	5	16.7	396	2	B75290	hypothetical prote
501	5	16.7	397	2	T00914	leucine-rich repa
502	5	16.7	398	2	S35264	cell division prot
503	5	16.7	398	2	D86700	hypothetical prote
504	5	16.7	403	2	T26326	hypothetical prote
505	5	16.7	404	2	G83322	hypothetical prote
506	5	16.7	406	2	T36196	probable acyl carr
507	5	16.7	409	2	T43599	yop targeted effec
508	5	16.7	411	2	S41648	translation elonga
509	5	16.7	411	2	S11961	hypothetical prote
510	5	16.7	412	1	OZ2QAF	circumsporozite p
511	5	16.7	412	1	I39554	phosphoglycerate k
512	5	16.7	412	2	T23385	hypothetical prote
513	5	16.7	413	2	I39551	phosphoglycerate k
514	5	16.7	413	2	T48635	hypothetical prote
515	5	16.7	414	2	S41609	acrolysin C (Ec 3.
516	5	16.7	414	2	AE1249	competence-damage
517	5	16.7	415	2	A11611	hypothetical prote
518	5	16.7	415	2	S48253	hypothetical prote
519	5	16.7	416	1	JN0006	nerve growth facto
520	5	16.7	416	2	T23097	hypothetical prote
521	5	16.7	417	2	S58193	adenosylhomocyste
522	5	16.7	419	1	DEECNB	glycerol-3-phospha
523	5	16.7	419	2	G91019	hypothetical prote
524	5	16.7	419	2	A85864	hypothetical prote
525	5	16.7	419	2	AC0792	glycerol-3-phospha
526	5	16.7	420	2	H86715	o-succinylbenzoate
527	5	16.7	421	2	C96553	hypothetical prote
528	5	16.7	421	2	T47393	hypothetical prote
529	5	16.7	421	2	S66714	probable membrane
530	5	16.7	422	2	A99285	hypothetical prote
531	5	16.7	423	2	H75058	hypothetical prote
532	5	16.7	424	2	A54533	circumsporozite p
533	5	16.7	424	2	C83902	maltose/maltodextr
534	5	16.7	424	2	F69723	trigger factor lig
535	5	16.7	424	1	A26431	nerve growth facto
536	5	16.7	425	1	KRXL2A	keratin, 64k type
537	5	16.7	425	1	AF2284	transcription term
538	5	16.7	426	2	B36467	hypothetical prote
539	5	16.7	427	1	T49031	acid phosphatase (
540	5	16.7	427	1	GOHUN	nerve growth facto

541	5	16.7	427	2	G86693	614	5	16.7	460	2	T11309	NADH2 dehydrogenas
542	5	16.7	428	2	AG1936	615	5	16.7	460	2	T110995	NADH2 dehydrogenas
543	5	16.7	428	2	B64301	616	5	16.7	461	1	QXLL4M	NADH2 dehydrogenas
544	5	16.7	428	2	S17817	617	5	16.7	462	2	S54572	probable phosphono
545	5	16.7	429	2	AC1629	618	5	16.7	463	2	B72500	bacteriochlorophyll
546	5	16.7	429	2	AI1266	619	5	16.7	463	2	D85943	conserved GTP bind
547	5	16.7	429	2	T45040	620	5	16.7	464	2	T29896	conserved GTP bind
548	5	16.7	431	2	E84031	621	5	16.7	465	2	T29896	hypothetical prote
549	5	16.7	432	2	T06851	622	5	16.7	466	2	T41375	hypothetical prote
550	5	16.7	432	2	G69977	623	5	16.7	466	2	C86739	probable phosphogl
551	5	16.7	433	2	S46668	624	5	16.7	467	2	B96979	hypothetical prote
552	5	16.7	433	2	S74922	625	5	16.7	468	2	A54926	spore germination
553	5	16.7	434	2	H97028	626	5	16.7	468	2	T49117	UDP-glucose 6-dehyd
554	5	16.7	435	2	AE1269	627	5	16.7	468	2	A82763	glucosidase like p
555	5	16.7	435	2	G70509	628	5	16.7	469	1	P2MLB	UDP-N-acetylmuram
556	5	16.7	436	2	S69821	629	5	16.7	469	2	D86309	l2 protein - bovin
557	5	16.7	436	2	T46107	630	5	16.7	469	2	B90110	hypothetical prote
558	5	16.7	436	2	B84845	631	5	16.7	471	2	A39024	TATA box-binding p
559	5	16.7	438	2	A72430	632	5	16.7	471	2	S76021	collagen alpha 3(I
560	5	16.7	439	2	D70958	633	5	16.7	472	2	T48074	probable trigger f
561	5	16.7	440	2	G71939	634	5	16.7	473	2	G72205	hypothetical prote
562	5	16.7	440	2	JC7807	635	5	16.7	473	2	T50920	hypothetical prote
563	5	16.7	441	2	C69834	636	5	16.7	475	1	CZCICA	cellulase (EC 3.2.
564	5	16.7	442	2	T07038	637	5	16.7	477	2	AC0122	cellulase (EC 3.2.
565	5	16.7	442	2	C90224	638	5	16.7	477	2	T47753	Sodium,galactoside
566	5	16.7	442	2	G71930	639	5	16.7	478	2	T10030	hypothetical prote
567	5	16.7	442	2	H91097	640	5	16.7	479	1	A32290	hypothetical prote
568	5	16.7	442	2	H89770	641	5	16.7	480	2	S76206	protein-tyrosine-p
569	5	16.7	443	2	T38239	642	5	16.7	481	2	B84700	hypothetical prote
570	5	16.7	443	2	B95229	643	5	16.7	481	2	A46602	probable flavonol
571	5	16.7	443	2	B84449	644	5	16.7	482	2	A46602	glutamate synthase
572	5	16.7	443	2	H98093	645	5	16.7	482	1	D83024	triacylglycerol 11
573	5	16.7	444	1	S75761	646	5	16.7	482	2	D83024	probable secretion
574	5	16.7	445	1	S03827	647	5	16.7	482	2	H70463	glycine dehydrogen
575	5	16.7	445	2	T30604	648	5	16.7	482	2	H84539	probable hemodoma
576	5	16.7	447	2	B81519	649	5	16.7	483	2	T15829	hypothetical prote
577	5	16.7	447	2	F72021	650	5	16.7	483	2	D71125	glutaryl-tRNA amid
578	5	16.7	447	2	A86603	651	5	16.7	483	2	S74637	lysine decarboxyla
579	5	16.7	447	2	S25817	652	5	16.7	484	2	JU0091	sucrose alpha-gluc
580	5	16.7	448	2	C64437	653	5	16.7	484	2	G72395	alpha-L-arabinofur
581	5	16.7	448	2	A43304	654	5	16.7	485	2	F82718	tldd protein XP112
582	5	16.7	449	2	B83086	655	5	16.7	485	2	T25199	hypothetical prote
583	5	16.7	449	2	B75451	656	5	16.7	486	2	AC2424	hypothetical prote
584	5	16.7	449	2	C69079	657	5	16.7	487	2	B84472	probable serine ca
585	5	16.7	449	2	B86763	658	5	16.7	487	2	S54265	glycoprotein gc -
586	5	16.7	451	1	S49016	659	5	16.7	487	2	H82298	D-alanyl-D-alanine
587	5	16.7	452	1	ALMSDS	660	5	16.7	488	2	T02257	hypothetical cell
588	5	16.7	452	2	T10882	661	5	16.7	488	2	T10675	nuclear antigen 21
589	5	16.7	452	2	A70389	662	5	16.7	489	2	D72523	hypothetical prote
590	5	16.7	454	2	H85293	663	5	16.7	489	2	E87518	hypothetical prote
591	5	16.7	455	2	B90619	664	5	16.7	490	2	T26983	magnesium transpor
592	5	16.7	455	2	T05804	665	5	16.7	491	2	S74473	hypothetical prote
593	5	16.7	455	2	G96708	666	5	16.7	491	2	B75554	probable starch sy
594	5	16.7	458	2	T11528	667	5	16.7	491	2	H89786	hypothetical prote
595	5	16.7	458	2	T11116	668	5	16.7	492	2	S71245	glucose-6-phosphat
596	5	16.7	458	2	B90615	669	5	16.7	492	2	F84602	hypothetical prote
597	5	16.7	458	2	B90617	670	5	16.7	492	2	C83592	hypothetical prote
598	5	16.7	458	2	B90625	671	5	16.7	494	2	JR0353	uridine diphosphog
599	5	16.7	458	2	B99613	672	5	16.7	497	2	G86299	F309.24 protein -
600	5	16.7	458	2	F81904	673	5	16.7	497	2	G81284	probable Ni/Fe-hyd
601	5	16.7	458	2	B75413	674	5	16.7	498	2	T09021	beta-glucosidase h
602	5	16.7	459	1	OXMS4M	675	5	16.7	499	2	E86182	hypothetical prote
603	5	16.7	459	1	T11463	676	5	16.7	505	2	B90181	Na+/H+ antiporter
604	5	16.7	459	2	T11084	677	5	16.7	507	2	B42249	serine-type carbox
605	5	16.7	459	2	T11398	678	5	16.7	508	2	S54264	glycoprotein gc -
606	5	16.7	459	2	T11398	679	5	16.7	508	2	G84564	probable sugar tra
607	5	16.7	459	2	T11489	680	5	16.7	509	2	A57512	beta-glucosidase B
608	5	16.7	459	2	T11450	681	5	16.7	509	2	T32201	hypothetical prote
609	5	16.7	459	2	S10196	682	5	16.7	510	2	F86392	Tik7.7 protein - A
610	5	16.7	459	2	S04756	683	5	16.7	510	2	P90367	n-methyl hydantoin
611	5	16.7	460	2	T11773	684	5	16.7	510	2	AI2012	hypothetical prote
612	5	16.7	460	2	T11296	685	5	16.7	511	1	S54720	glucose-6-phosphat
613	5	16.7	460	2	T11543	686	5	16.7	511	2	B82941	ATP synthase alpha

687	5	16.7	511	2	E84685	probable seed stor	760	5	16.7	563	1	FN0493	triacylglycerol 1i
688	5	16.7	513	2	S21535	nitrate reductase	761	5	16.7	563	2	B70918	hypothetical prote
689	5	16.7	513	2	JC5920	potassium channel	762	5	16.7	565	2	AF1325	alpha-acetolactate
690	5	16.7	513	2	C66897	hypothetical prote	763	5	16.7	565	2	H72253	DNA primase - Ther
691	5	16.7	513	2	E72741	hypothetical prote	764	5	16.7	565	2	T47330	hypothetical prote
692	5	16.7	514	2	AG0474	threonine ammonia-	765	5	16.7	565	2	S27495	nodu protein - Bra
693	5	16.7	515	1	O4H04	cytochrome P450 1A	766	5	16.7	570	2	E97155	DNA repair protein
694	5	16.7	517	1	S28229	diacylglycerol kin	767	5	16.7	570	2	S24459	hypothetical prote
695	5	16.7	520	2	E90304	conserved hypochet	768	5	16.7	572	2	T51575	2-hydroxyphenoyl
696	5	16.7	520	2	B97257	lyam domain conta	769	5	16.7	572-	2	H96685	probable AMP-bind
697	5	16.7	521	1	VCBEHB	glycoprotein gIII	770	5	16.7	573	2	T01317	probable pectinest
698	5	16.7	521	2	DB3399	probable ATP-bind	771	5	16.7	573	2	B97331	membrane associat
699	5	16.7	521	2	T21440	hypothetical prote	772	5	16.7	575	2	AG2586	hypothetical prote
700	5	16.7	523	2	S47053	hypothetical prote	773	5	16.7	577	2	T47838	beta-glucosidase-1
701	5	16.7	525	2	H70982	probable fadD7 pro	774	5	16.7	577	2	T02401	probable beta-gluc
702	5	16.7	525	2	T32481	hypothetical prote	775	5	16.7	577	2	T41727	F-box domain prote
703	5	16.7	525	2	T12685	hypothetical prote	776	5	16.7	580	2	T43732	Crp synthase (EC 6
704	5	16.7	527	2	S39549	thioglucoisidase (E	777	5	16.7	580	2	A96683	hypothetical prote
705	5	16.7	527	2	G86390	TIK7.22 protein -	778	5	16.7	580	2	B84552	similar to Mlo pro
706	5	16.7	528	2	B64760	propionate catabol	779	5	16.7	582	2	A55174	kinase-associated
707	5	16.7	528	2	C85527	regulator for prp	780	5	16.7	582	2	I50224	erythroid cell tra
708	5	16.7	528	2	H90676	regulator for prp	781	5	16.7	583	2	E70762	probable ftda prot
709	5	16.7	528	2	S32593	[RNA-polymerase]-s	782	5	16.7	583	2	B85063	hypothetical prote
710	5	16.7	528	2	T33464	hypothetical prote	783	5	16.7	583	2	F97492	probable transport
711	5	16.7	530	2	B86326	T29M8.12 protein -	784	5	16.7	586	2	H86914	conserved hypochet
712	5	16.7	531	1	A39740	sterol 27-monooxyg	785	5	16.7	586	2	I37202	B-CAM protein - hu
713	5	16.7	531	2	T43551	multidrug resistanc	786	5	16.7	593	2	D97806	penicillin-binding
714	5	16.7	532	2	A35149	lpnH protein - Shi	787	5	16.7	594	2	C71661	hypothetical prote
715	5	16.7	532	2	T05281	probable homeobox	788	5	16.7	597	2	T45676	hypothetical prote
716	5	16.7	532	2	AE1964	Dnak-type molecula	789	5	16.7	599	2	S18735	centromere protein
717	5	16.7	533	2	T40011	hypothetical prote	790	5	16.7	600	2	E86854	hypothetical prote
718	5	16.7	535	1	A33813	sterol 26-monooxyg	791	5	16.7	601	2	T18787	hypothetical prote
719	5	16.7	535	2	F70914	probable acid-COA	792	5	16.7	602	2	C82581	GTP binding protei
720	5	16.7	535	2	C96706	probable amp-bind	793	5	16.7	605	2	T40556	hypothetical prote
721	5	16.7	535	2	C86772	conserved hypochet	794	5	16.7	605	2	T31557	hypothetical prote
722	5	16.7	535	2	AE2710	MS permease [impo	795	5	16.7	606	2	S70358	centromere protein
723	5	16.7	535	2	AD2107	serine/threonine k	796	5	16.7	607	2	T20796	hypothetical prote
724	5	16.7	536	2	S56654	thioglucoisidase (E	797	5	16.7	609	2	T40660	hypothetical C2H2-
725	5	16.7	537	2	S25181	heat shock protein	798	5	16.7	611	2	T41563	hypothetical prote
726	5	16.7	538	2	S62566	PHF finger protein	799	5	16.7	611	2	T03890	hypothetical prote
727	5	16.7	539	2	T51215	hypothetical prote	800	5	16.7	611	2	C84863	hypothetical prote
728	5	16.7	540	2	B96747	probable alanine a	801	5	16.7	613	2	S15468	complement C3b/C4b
729	5	16.7	540	2	B83121	probable AMP-bind	802	5	16.7	613	2	S75976	hypothetical prote
730	5	16.7	541	2	S56653	thioglucoisidase (E	803	5	16.7	614	2	T20795	hypothetical prote
731	5	16.7	541	2	AF0547	propionate catabol	804	5	16.7	614	2	T47749	hypothetical prote
732	5	16.7	541	2	H83445	probable chemotaxi	805	5	16.7	616	2	B84500	probable retroelem
733	5	16.7	543	2	S46713	ATP-dependent RNA	806	5	16.7	617	2	T02121	hypothetical prote
734	5	16.7	544	2	S41094	triacylglycerol 1i	807	5	16.7	617	2	G82706	conserved hypochet
735	5	16.7	544	2	S41095	triacylglycerol 1i	808	5	16.7	619	2	F82984	hypothetical prote
736	5	16.7	544	2	S41096	triacylglycerol 1i	809	5	16.7	621	2	A75101	aldelyde-ferredoxi
737	5	16.7	544	2	S19149	thioglucoisidase (E	810	5	16.7	624	2	T04414	probable glucan 1,
738	5	16.7	545	2	T02279	hypothetical prote	811	5	16.7	625	2	G71072	aldelyde-ferredoxi
739	5	16.7	545	2	A87136	hypothetical prote	812	5	16.7	625	2	B97358	glucose-inhibited
740	5	16.7	546	1	S71008	propionyl-CoA carb	813	5	16.7	627	2	S40048	1,4-alpha-glucan b
741	5	16.7	546	2	G86221	protein F7G19.14 [	814	5	16.7	627	2	T02846	dynein light chain
742	5	16.7	546	2	C69450	anthranilate synh	815	5	16.7	627	2	G84558	probable SEC1 fami
743	5	16.7	547	2	S56656	thioglucoisidase (E	816	5	16.7	628	2	I38000	Lutheran blood gro
744	5	16.7	548	2	S26149	thioglucoisidase (E	817	5	16.7	629	2	C82598	glucose inhibited
745	5	16.7	548	2	C82698	electron transfer	818	5	16.7	634	2	C70133	conserved hypochet
746	5	16.7	550	2	S55118	probable membrane	819	5	16.7	636	2	H96666	probable phytochro
747	5	16.7	551	2	F75045	dihydroxy-acid deh	820	5	16.7	638	2	T47569	hypothetical prote
748	5	16.7	551	2	H96682	hypothetical prote	821	5	16.7	640	2	S47865	regulatory protein
749	5	16.7	553	2	C71257	phenylalanine-tRNA	822	5	16.7	642	2	C81345	hypothetical prote
750	5	16.7	553	2	S70817	invasion genes tra	823	5	16.7	645	2	S49570	penicillin-binding
751	5	16.7	553	2	AH0849	invasion protein r	824	5	16.7	645	2	G81315	membrane bound zin
752	5	16.7	555	2	B72486	probable hydanotin	825	5	16.7	645	2	T05251	probable disease x
753	5	16.7	558	2	A43667	serine C-palmitoyl	826	5	16.7	647	2	F90595	conserved hypochet
754	5	16.7	559	2	T22179	hypothetical prote	827	5	16.7	648	2	JC4674	Slly1 protein - rat
755	5	16.7	559	2	P96912	adenylylulafate re	828	5	16.7	649	2	A33589	mismatch repair pr
756	5	16.7	561	1	S52319	unspecific monooxy	829	5	16.7	649	2	A95020	DNA mismatch repai
757	5	16.7	561	2	C64459	dihydroxy-acid den	830	5	16.7	649	2	H97891	DNA mismatch repai
758	5	16.7	562	2	S55516	glucan 1,3-beta-gl	831	5	16.7	654	2	T21618	hypothetical prote
759	5	16.7	562	2	S38149	SIS2 protein - yea	832	5	16.7	655	1	S50096	probable splicing



833	5	16.7	663	2	T30621	hypothetical prote
834	5	16.7	664	2	T35122	probable ATP-depen
835	5	16.7	665	2	D96621	auxin response fac
836	5	16.7	667	2	T33526	hypothetical prote
837	5	16.7	668	2	S64123	hypothetical prote
838	5	16.7	671	2	B75607	conserved hypotet
839	5	16.7	672	2	G88651	protein B0212.3 [l
840	5	16.7	678	2	T42668	hypothetical prote
841	5	16.7	679	2	S48939	hypothetical prote
842	5	16.7	684	2	G70744	hypothetical prote
843	5	16.7	685	2	S64985	hypothetical prote
844	5	16.7	688	2	S39491	hypothetical prote
845	5	16.7	688	2	S32961	protochlorophyllid
846	5	16.7	690	2	A12516	hypothetical prote
847	5	16.7	698	2	C69208	DNA helicase relat
848	5	16.7	691	2	T25519	hypothetical prote
849	5	16.7	695	2	H86900	DNA mismatch repai
850	5	16.7	697	2	T49727	related to long-ch
851	5	16.7	698	2	C84836	hypothetical prote
852	5	16.7	699	2	G86311	hypothetical prote
853	5	16.7	702	2	D86469	protein F12K21.12
854	5	16.7	705	2	S45769	probable membrane
855	5	16.7	714	2	T25744	hypothetical prote
856	5	16.7	715	2	H90977	hypothetical prote
857	5	16.7	715	2	B75135	DNA helicase relat
858	5	16.7	716	2	AG2446	single-stranded-DN
859	5	16.7	719	2	F85824	hypothetical prote
860	5	16.7	720	2	T43854	helicase (imported
861	5	16.7	721	2	C82939	virulence associat
862	5	16.7	723	2	C84507	hypothetical prote
863	5	16.7	725	2	T52158	hypothetical prote
864	5	16.7	728	2	A75582	serine proteinase,
865	5	16.7	728	2	S57142	hypothetical prote
866	5	16.7	731	2	D71332	probable DNA topoi
867	5	16.7	738	2	E86294	hypothetical prote
868	5	16.7	740	2	D97180	relA/SpoT protein,
869	5	16.7	741	2	T05250	probable disease r
870	5	16.7	743	2	T02147	hypothetical prote
871	5	16.7	746	2	AD1622	probable integral
872	5	16.7	749	2	H82691	topoisomerase IV s
873	5	16.7	754	1	PIV50	RNA-directed RNA p
874	5	16.7	755	2	T20950	hypothetical prote
875	5	16.7	756	2	T00367	hypothetical prote
876	5	16.7	761	1	B64506	DNA topoisomerase
877	5	16.7	762	2	T00410	protein kinase hom
878	5	16.7	765	2	S38099	hypothetical prote
879	5	16.7	766	2	T48389	cucumisin-like pro
880	5	16.7	767	2	S63182	hypothetical prote
881	5	16.7	768	1	A34106	protein kinase (EC
882	5	16.7	768	2	G82251	cation transport A
883	5	16.7	768	2	T17462	disease resistance
884	5	16.7	769	2	S50966	probable membrane
885	5	16.7	772	2	C69980	transcription regu
886	5	16.7	772	2	B82888	phenylalanine-UNA
887	5	16.7	774	2	D83208	probable ferredoxi
888	5	16.7	774	2	S25284	protein kinase nek
889	5	16.7	779	2	E71825	probable type III
890	5	16.7	780	2	AB1801	amino-terminal dom
891	5	16.7	780	2	AC1427	transcription regu
892	5	16.7	781	2	A56244	DNA repair/recombi
893	5	16.7	782	2	T32155	hypothetical prote
894	5	16.7	785	2	S46672	hypothetical prote
895	5	16.7	788	1	QOBER3	HLH1 protein - hu
896	5	16.7	788	2	H96545	hypothetical prote
897	5	16.7	790	2	S61587	transcription acti
898	5	16.7	790	2	D86528	omp85 analog (impo
899	5	16.7	791	2	D72094	omp85 analog - Chl
900	5	16.7	791	2	A46140	diacylglycerol kin
901	5	16.7	792	2	T22717	hypothetical prote
902	5	16.7	795	2	T48252	ecceiferum3 (CER3)
903	5	16.7	795	2	AF2444	hypothetical prote
904	5	16.7	796	2	B46140	diacylglycerol kin
905	5	16.7	796	2	JC7355	peroxisome prolife
906	5	16.7	802	2	T32448	hypothetical prote
907	5	16.7	803	2	S45582	siderophore biosyn
908	5	16.7	809	2	T16448	hypothetical prote
909	5	16.7	811	2	T17863	vitellogenin II ho
910	5	16.7	813	2	G75331	ATP-dependent prot
911	5	16.7	815	2	T76775	09 mannan biosynth
912	5	16.7	815	2	AB2444	hypothetical prote
913	5	16.7	816	2	B95093	hypothetical prote
914	5	16.7	816	2	C88196	protein ZK1127.7 [
915	5	16.7	816	2	G97960	ATP-dependent DNA
916	5	16.7	818	2	F97772	hypothetical prote
917	5	16.7	818	2	T15380	hypothetical prote
918	5	16.7	819	2	B47057	aspartate kinase (
919	5	16.7	821	2	S67087	hypothetical prote
920	5	16.7	829	2	S75776	pled protein - Syn
921	5	16.7	830	1	RNEGB2	DNA-directed RNA p
922	5	16.7	831	2	B40204	Na+/H+-exchanging
923	5	16.7	832	2	A40205	Na+/H+-exchanging
924	5	16.7	833	1	A31593	heat shock transcr
925	5	16.7	835	2	E71691	outer membrane ass
926	5	16.7	835	2	T05259	probable disease r
927	5	16.7	836	1	QOVZRA	mRNA guanylyltrans
928	5	16.7	838	2	A32262	fatty-acid synthas
929	5	16.7	839	2	H97758	outer membrane ass
930	5	16.7	840	2	H89911	oxacillin resistanc
931	5	16.7	840	2	S48975	hypothetical prote
932	5	16.7	843	2	F96537	hypothetical prote
933	5	16.7	846	1	QOBER3	HORF1 protein - hu
934	5	16.7	849	2	T14519	probable S-recepto
935	5	16.7	850	2	AB1954	hypothetical prote
936	5	16.7	851	2	S33985	env polypeptid - hu
937	5	16.7	854	2	S13288	env polypeptid - hu
938	5	16.7	856	1	VCLJVL	env polypeptid pr
939	5	16.7	856	1	VCLJVL	env polypeptid pr
940	5	16.7	856	2	C85023	hypothetical prote
941	5	16.7	859	2	F69159	protoporphyrin IX
942	5	16.7	860	2	C86203	hypothetical prote
943	5	16.7	861	1	VCLJLV	env polypeptid pr
944	5	16.7	861	2	H64102	leucine-tRNA ligas
945	5	16.7	863	2	G96964	probable permease,
946	5	16.7	864	2	T08875	protein kinase hom
947	5	16.7	864	2	H85335	hypothetical prote
948	5	16.7	864	2	T04518	hypothetical prote
949	5	16.7	867	2	T27136	hypothetical prote
950	5	16.7	867	2	T14777	hypothetical prote
951	5	16.7	870	2	G81348	valine-tRNA ligase
952	5	16.7	870	2	B71698	hypothetical prote
953	5	16.7	871	2	T27135	hypothetical prote
954	5	16.7	873	2	F86426	95.1K hypothetical
955	5	16.7	874	2	AC2287	hypothetical prote-
956	5	16.7	877	2	T43449	hypothetical prote
957	5	16.7	878	2	B84977	alanine-tRNA ligas
958	5	16.7	879	2	S49910	chloroplast outer
959	5	16.7	882	2	B96931	hypothetical prote
960	5	16.7	889	2	H96606	hypothetical prote
961	5	16.7	892	2	J00424	probable transposa
962	5	16.7	892	2	T01899	disease resistance
963	5	16.7	895	1	EXCKP	H+-exporting ATPas
964	5	16.7	895	2	B96775	hypothetical prote
965	5	16.7	896	2	T07408	lipoygenase (EC 1
966	5	16.7	896	2	J02391	lipoygenase (EC 1
967	5	16.7	898	2	T21179	hypothetical prote
968	5	16.7	899	2	T07062	probable lipoxigen
969	5	16.7	900	2	C71339	probable DNA misma
970	5	16.7	904	2	C70559	probable po1a prot
971	5	16.7	905	2	T39572	probable proteinas
972	5	16.7	908	2	A83424	asacetylatory nitra
973	5	16.7	910	2	B83451	aconitinate hydrat
974	5	16.7	911	2	F81974	leucine-tRNA ligas
975	5	16.7	911	2	S77659	DNA-directed DNA p
976	5	16.7	911	2	A39967	inter-alpha-trypsi
977	5	16.7	917	2	T04661	hypothetical prote
978	5	16.7	918	2	AB2445	hypothetical prote

979	5	16.7	920	2	T41050	conserved hypothet
980	5	16.7	924	2	T25007	hypothetical prote
981	5	16.7	925	2	JC2033	G protein-coupled
982	5	16.7	925	2	T16235	hypothetical prote
983	5	16.7	928	1	S38001	probable serine/ch
984	5	16.7	928	1	AC1312	ATP-dependent heli
985	5	16.7	928	2	AC1684	ATP-dependent heli
986	5	16.7	930	2	AC2412	preprotein translo
987	5	16.7	932	2	T25008	hypothetical prote
988	5	16.7	932	2	H69045	hypothetical prote
989	5	16.7	936	2	T06190	lipoxigenase (EC 1
990	5	16.7	941	2	T33032	hypothetical prote
991	5	16.7	942	2	G83038	probable sensor/re
992	5	16.7	943	2	T24707	hypothetical prote
993	5	16.7	944	2	A89624	protein F21A10.2 (
994	5	16.7	947	2	H85088	hypothetical prote
995	5	16.7	950	2	BE4135	oxoglutarate dehyd
996	5	16.7	951	1	STECVT	valine-tRNA ligase
997	5	16.7	951	2	C91283	valine tRNA synthe
998	5	16.7	951	2	E86124	valine-tRNA ligase
999	5	16.7	951	2	AC1061	valine-tRNA ligase
1000	5	16.7	956	2	B71468	probable insulinase

## ALIGNMENTS

## RESULT 1

DSBYN

superoxide dismutase (EC 1.15.1.1) (Mn) precursor [validated] - Yeast (Saccharomyces cer

N;Alternate names: protein YHR008C

C;Species: Saccharomyces cerevisiae

C;Date: 19-Feb-1994 #sequence revision 17-Mar-1987 #text\_change 23-Mar-2001

C;Accession: A00521; S46785; A90766

R;Marrez, C.A.M.; Van Loon, A.P.G.M.; Oudhoorn, P.; Van Steeg, H.; Grivell, L.A.; Slate

Eur. J. Biochem. 147, 153-161, 1985

A;Title: Nucleotide sequence analysis of the nuclear gene coding for manganese superoxid

A;Reference number: A91141; MUID:85127011; PMID:3882422

A;Accession: A00521

A;Molecule type: DNA

A;Residues: 1-233 &lt;MAR&gt;

A;Cross-references: EMBL:X02156; NID:G4513; PIDN:CAA26092.1; PID:G4514

R;Du, Z.

A;Submitted to the EMBL Data Library, June 1994

A;Description: The sequence of S. cerevisiae cosmid L2825.

A;Reference number: S46774

A;Accession: S46785

A;Molecule type: DNA

A;Residues: 1-233 &lt;DUZ&gt;

A;Cross-references: EMBL:U10400; NID:G500701; PIDN:AB68939.1; PID:G500704; GSPDB:GN0000

K;Dittlow, C.; Johansen, J.T.; Martin, B.M.; Svendsen, I.

Carlsberg Res. Commun. 47, 81-91, 1982

A;Title: The complete amino acid sequence of manganese-superoxide dismutase from Sacchar

A;Reference number: A90766

A;Accession: A90766

A;Molecule type: Protein

A;Residues: 27-229 &lt;DT&gt;

C;Genetics:

A;Gene: SOD2, MIP8:YHR008C

A;Cross-references: SGD:S0001050; MIP8:YHR008C

A;Map position: 8R

A;Genome: nuclear

C;Complex: nucleotriamer

C;Function:

A;Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen

C;Superfamily: superoxide dismutase (Mn)

C;Keywords: homotrimer; manganese; metalloprotein; mitochondrial matrix; mitochondrion

F;1-26/Domain: transist peptide (mitochondrion) #status predicted &lt;TMP&gt;

F;27-229/Product: superoxide dismutase (Mn) #status experimental &lt;MNT&gt;

F;52,107,194,198/Binding site: manganese (His, His, Asp, His) #status predicted

Query Match 33.3%; Score 10; DB 1; Length 233;

Best local similarity 100.0%; Pred. No. 0.002;

Matches	10;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	15	EPYISQINE 24							
Db	41	EPYISQINE 50							

## RESULT 2

superoxide dismutase (EC 1.15.1.1) (Mn) - Nocardia asteroides

C;Species: Nocardia asteroides

C;Date: 22-Dec-1995 #sequence revision 08-Feb-1996 #text\_change 05-May-2000

C;Accession: J04351

R;Alcendor, D.J.; Chapman, G.D.; Beaman, B.L.

Gene 164, 143-147, 1995

A;Title: Isolation, sequencing and expression of the superoxide dismutase-encoding gene

A;Reference number: J04351; MUID:96060854; PMID:7550304

A;Accession: J04351

A;Molecule type: DNA

A;Residues: 1-207 &lt;ALC&gt;

A;Cross-references: GB:U02341

A;Experimental source: GUN-2

A;Note: The authors translated the initiation codon GTG for residue 1 as Val

C;Genetics:

A;Gene: sod

A;Start codon: GTG

C;Function:

A;Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen

C;Keywords: manganese; metalloprotein; oxidoreductase; virulence factor

F;28,76,160,164/Binding site: manganese (His, His, Asp, His) #status predicted

Query Match 23.3%; Score 7; DB 2; Length 207;

Best local similarity 100.0%; Pred. No. 2.9;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 ISQINE 24

|||||

Db 20 ISQINE 26

## RESULT 3

superoxide dismutase (EC 1.15.1.1) (Fe) [validated] - Mycobacterium tuberculosis

C;Species: Mycobacterium tuberculosis

C;Date: 21-Nov-1993 #sequence revision 26-May-1995 #text\_change 20-Jun-2000

C;Accession: S15205; C70654; S10908

R;Zhang, Y.; Lathigra, R.; Garbe, T.; Catty, D.; Young, D.

Mol. Microbiol. 5, 381-391, 1991

A;Title: Genetic analysis of superoxide dismutase, the 23 kilodalton antigen of Mycobac

A;Reference number: S15205; MUID:91251768; PMID:1904126

A;Accession: S15205

A;Molecule type: DNA

A;Residues: 1-207 &lt;ZHA&gt;

A;Cross-references: EMBL:X52861; NID:G794079; PIDN:CAA37042.1; PID:G581379

A;Experimental source: strain H37Rv

A;Note: the authors translated the codon GAA for residue 3 as Gln

R;Cole, S.T.; Broesch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon,

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrett, B.G.

A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A;Reference number: A70500; MUID:98295987; PMID:9634230

A;Accession: C70654

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-207 &lt;COL&gt;

A;Cross-references: GB:Z83864; GB:AL123456; NID:G3261687; PIDN:CAB06220.1; PID:G1781110

A;Experimental source: strain H37Rv

R;Cooper, J.B.; McInyre, K.; Wood, S.P.; Zhang, Y.; Young, D.

submitted to the Brookhaven Protein Data Bank, September 1994

A;Reference number: A52807; PDB:1IDS

A;Contents: annotation; X-ray crystallography, 2.0 angstroms, residues 2-199  
 R;Cooper, J.B.; Driessen, H.P.C.; Wood, S.P.; Zhang, Y.; Young, D.  
 J. Mol. Biol. 235, 1156-1158, 1994  
 A;Title: Crystallisation and preliminary X-ray analysis of the iron-dependent superoxide  
 A;Reference number: A59029; MUID:94118350; PMID:8289318  
 A;Contents: annotation; X-ray crystallography  
 C;Genetics:  
 A;Gene: soga  
 A;Start codon: GTG  
 C;Function:  
 A;Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen  
 C;Superfamily: superoxide dismutase (Mn)  
 C;Keywords: iron; metalloprotein; oxidoreductase; tetramer  
 F;28,76,160,164/Binding site: iron (His, His, Asp, His) #status experimental

Query Match 23.3%; Score 7; DB 2; Length 207;  
 Best Local Similarity 100.0%; Pred. No. 2.9;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

18 ISGQINE 24  
 |||||  
 20 ISGQINE 26

RESULT 4  
 S60669  
 superoxide dismutase (EC 1.15.1.1) (Mn) - Mycobacterium fortuitum  
 C;Species: Mycobacterium fortuitum  
 C;Date: 19-Jul-1996 #sequence\_revision 13-Mar-1997 #text\_change 18-Jun-1999  
 C;Accession: S60669  
 R;Kendrick, C.; Domenech, P.; Prieto, J.; Garcia, M.J.  
 submitted to the EMBL Data Library, August 1995  
 A;Description: Cloning and expression of the superoxide dismutase gene of Mycobacterium  
 A;Reference number: S60669  
 A;Accession: S60669  
 A;Molecule type: DNA  
 A;Residues: 1-207 <MEN>  
 A;Cross-references: EMBL:X70914; NID:9550078; PIDN:CA550266.1; PID:9550079  
 A;Experimental source: ATCC 6841  
 C;Genetics:  
 A;Gene: sod  
 A;Start codon: GTG  
 C;Function:  
 A;Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen  
 C;Superfamily: superoxide dismutase (Mn)  
 C;Keywords: manganese; metalloprotein; oxidoreductase  
 F;28,76,160,164/Binding site: manganese (His, His, Asp, His) #status predicted

Query Match 23.3%; Score 7; DB 2; Length 207;  
 Best Local Similarity 100.0%; Pred. No. 2.9;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 ISGQINE 24  
 |||||  
 DB 20 ISGQINE 26

RESULT 5  
 T03296  
 beta-glucosidase (EC 3.2.1.21), chloroplast - rice  
 N;Alternate names: beta-D-glucoside glucosylhydrolase  
 C;Species: Oryza sativa (rice)  
 C;Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 24-Mar-1999  
 C;Accession: T03296  
 R;Ben, A.  
 submitted to the EMBL Data Library, May 1995  
 A;Reference number: Z14892  
 A;Accession: T03296  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-328 <ESE>  
 A;Cross-references: EMBL:U028047; NID:G1143863; PID:G1143864  
 C;Genetics:

A;Genome: nuclear  
 C;Function:  
 A;Description: catalyzes the release of either gibberellin or cyanogenic substances from  
 C;Keywords: chloroplast; glucosidase; hydrolase

Query Match 23.3%; Score 7; DB 2; Length 328;  
 Best Local Similarity 100.0%; Pred. No. 4.5;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 SATPEYI 18  
 |||||  
 DB 56 SATPEYI 62

RESULT 6  
 B69959  
 glycine dehydrogenase homolog yghK - Bacillus subtilis  
 C;Species: Bacillus subtilis  
 C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 20-Jun-2000  
 C;Accession: B69959  
 R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte  
 C.; Bron, S.; Brouillet, S.; Brunsch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch  
 A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferreri, E.  
 Nature 390, 249-256, 1997  
 A;Authors: Foulger, D.; Filtz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galle  
 lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holzapfel, S.; Hosono, S.; Hullo, M.F  
 Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois  
 Y.; Muthers, lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Meuse  
 A.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Paro, V.; Pohl, T.M.; Portetelli  
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon  
 A;Authors: Schleich, S.; Schreiber, R.; Scottone, F.; Sekiguchi, J.; Sekowska, A.; Serc  
 akenchi, M.; Tamakoshi, A.; Tanaka, T.; Terstira, P.; Tognoni, A.; Tosato, V.; Uchiyama  
 T.; Winkler, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,  
 A;Authors: Yoshikawa, H.F.; Zunshtein, E.; Yoshikawa, H.; Danchin, A.  
 A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
 A;Reference number: A69580; MUID:98044033; PMID:9384377  
 A;Accession: B69959  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-488 <KUN>  
 A;Cross-references: GB:Z99116; GB:AL009126; NID:G2634723; PIDN:CA14386.1; PID:G2634889  
 A;Experimental source: strain 168  
 C;Genetics:  
 A;Gene: yghK

Query Match 23.3%; Score 7; DB 2; Length 488;  
 Best Local Similarity 100.0%; Pred. No. 6.6;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YSLPELD 8  
 |||||  
 DB 20 YSLPELD 26

RESULT 7  
 T10791  
 beta-glucosidase (EC 3.2.1.21) - cassava  
 N;Alternate names: linamarase  
 C;Species: Manihot esculenta (cassava)  
 C;Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 16-Jul-1999  
 C;Accession: T10791  
 R;McMahon, J.M.; Sayre, R.T.  
 submitted to the EMBL Data Library, March 1997  
 A;Description: Genomic sequence for a linamarase gene from cassava (Manihot esculenta C  
 A;Reference number: Z17147  
 A;Accession: T10791  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-507 <MCM>  
 A;Cross-references: EMBL:U05298; NID:G2443453; PID:G2443454  
 A;Experimental source: strain MTail  
 C;Genetics:  
 A;Inserts: 27/2; 48/1; 67/3; 93/1; 119/1; 149/3; 233/3; 273/2; 344/3; 356/3; 391/1; 427

A>Note: PLIN-GEN  
C:Function:  
A:Description: responsible for hydrolysis of two structurally related cyanoglucosides (1  
C:Superfamily: Agrobacterium beta-glucosidase  
C:Keywords: glycosidase; hydrolase

Query Match 23.3%; Score 7; DB 2; Length 507;  
Best Local Similarity 100.0%; Pred. No. 6.9;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 12 SATEPYI 18  
Db 206 SATEPYI 212

RESULT 8  
S23940  
beta-glucosidase (EC 3.2.1.21) - cassava  
N:Alternate names: linamarase  
C:Species: Manihot esculenta (cassava)  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 22-Jun-1999  
Accession: S23940; S51196

R:Hughes, M.A.; Brown, K.; Pancoro, A.; Murray, B.S.; Oxtoby, E.; Hughes, J.  
Arch. Biochem. Biophys. 295, 273-279, 1992  
A:Title: A molecular and biochemical analysis of the structure of the cyanogenic beta-gl  
A:Reference number: S23940; PMID:92264724; PMID:1586156  
A:Accession: S23940  
A:Molecule type: mRNA  
A:Residues: 1-531 <HUG>  
A:Cross-references: GB:S51175; NID:G249261; PIDN:AMB22162.1; PID:G249262  
A:Experimental source: seed  
A:Keywords: Z.; Klass, L.; Hughes, M.A.  
Arch. Biochem. Biophys. 315, 323-330, 1994  
A:Title: Investigation of the active site of the cyanogenic beta-D-glucosidase (linamar  
with acid catalytic function.  
A:Reference number: S51196; PMID:95077406; PMID:7986074  
A:Accession: S51196  
A:Molecule type: Protein  
A:Residues: 128-215 <KER>  
C:Function:

A:Description: responsible for hydrolysis of two structurally related cyanoglucosides (1  
C:Superfamily: Agrobacterium beta-glucosidase  
C:Keywords: glycoprotein; glycosidase; hydrolase  
F:198/Active site: Glu #status experimental  
F:366,398,421,489,506/Binding site: carbohydrate (Aan) (covalent) #status predicted

Query Match 23.3%; Score 7; DB 2; Length 531;  
Best Local Similarity 100.0%; Pred. No. 7.2;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 12 SATEPYI 18  
Db 230 SATEPYI 236

RESULT 9  
S76134  
hypothetical protein - Synecocystis sp. (strain PCC 6803)  
C:Species: Synecocystis sp.  
A:Variety: PCC 6803  
C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 20-Jun-2000

R:Keneko, T.; Saito, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;  
o, K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Matanabe, A.; Yamada, M.; Yasuda  
DNA Res. 3, 103-116, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis  
8.  
A:Reference number: S74332; PMID:97061201; PMID:8905231  
A:Accession: S76134  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1034 <KAN>  
A:Cross-references: EMBL:D90914; GB:AB001339; NID:G1653477; PIDN:BAA18393.1; PID:G165348

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
C:Superfamily: phosphoenolpyruvate carboxylase

Query Match 23.3%; Score 7; DB 2; Length 1034;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KYSLPEL 7  
Db 754 KYSLPEL 760

RESULT 10  
T31617  
hypothetical protein Y5088A.m - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
Accession: T31617  
R:Steward, C.  
submitted to the EMBL Data Library, September 1999  
A:Accession: T31617  
A:Reference number: Z21047

A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-2055 <WIL>  
A:Cross-references: EMBL:AL117200; NID:e1549770; PIDN:CAB55056.1; CESP:Y5088A.m  
A:Experimental source: clone Y5088A  
C:Genetics:  
A:Gene: CESP:Y5088A.m  
A:Introns: 273/3; 447/1; 526/1; 735/1; 1247/1; 1418/1; 1494/1; 1753/3

Query Match 23.3%; Score 7; DB 2; Length 2055;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 5 PELDYEF 11  
Db 1780 PELDYEF 1786

RESULT 11  
AH1065  
hypothetical protein STY4854 [imported] - Salmonella enterica subsp. enterica serovar T  
C:Species: Salmonella enterica subsp. enterica serovar Typh  
A>Note: this species has also been called Salmonella typh  
C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 09-Nov-2001  
Accession: AH1065  
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher  
ch, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar  
, S.; Moule, S.; O'Gaora, P.  
Nature 413, 848-852, 2001

A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skellon, J.; Stevens, K.  
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica sero  
A:Reference number: AB0502; PMID:11677608  
A:Accession: AH1065  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-91 <PAR>  
A:Cross-references: GB:AL513382; PIDN:CAD06573.1; PID:G16505614; GSPDB:GN00176  
C:Genetics:  
A:Gene: STY4854

Query Match 20.0%; Score 6; DB 2; Length 91;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 17 YISQOI 22  
Db 79 YISQOI 84

RESULT 12  
JC4396

superoxide dismutase (EC 1.15.1.1) (Fe/Mn) [validated] - *Propionibacterium freudenreichii*  
C/Species: *Propionibacterium freudenreichii* subsp. *shermanii*  
C/Date: 20-Jan-1996 #sequence\_revision 19-Apr-1996 #text\_change 20-Apr-2000  
C/Accession: J04396; S41106  
R/Gibbanielli, R.; Battistoni, A.; Polizio, F.; Carri, M.T.; De Martino, A.; Meier, B.; Biochem. Biophys. Res. Commun. 216, 841-847, 1995  
A/Title: Metal uptake of recombinant catalytic superoxide dismutase from *Propionibacterium*  
A/Reference number: J04396; MUID:96074560; PMID:7488202  
A/Accession: J04396  
A/Molecule type: DNA  
A/Residues: 1-202 <GB>  
A/Cross-references: EMBL:X01650  
A/Experimental source: PZ3  
R/Meier, B.; Selin, A.P.; Schinina, M.E.; Barra, D.  
Eur. J. Biochem. 219, 463-468, 1994  
A/Title: In vivo incorporation of copper into the iron-exchangeable and manganese-exchangeable sites.  
A/Reference number: S41106; MUID:94139724; PMID:8307013  
A/Accession: S41106  
A/Molecule type: protein  
A/Residues: 2-202 <MEI>  
A/Experimental source: strain PZ3  
C/Genetics:  
A/Gene: *sod*  
A/Complex: homotetramer  
C/Function:  
A/Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen  
A/Note: can use iron or manganese as cofactor  
C/Superfamily: superoxide dismutase (Mn)  
C/Keywords: homotetramer; iron; manganese; metalloprotein; oxidoreductase  
F/2-202/Procydot: superoxide dismutase #status experimental <MAT>  
F/28,76,162,166/Binding site: iron/manganese (His, His, Asp, His) #status predicted

Query Match 20.0%; Score 6; DB 2; Length 202;  
Best Local Similarity 100.0%; Pred. No. 34;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 EPIYISG 20  
|||||  
Db 17 EPIYISG 22

RESULT 13  
B86498  
Superoxide dismutase (Mn) [imported] - *Chlamydomonas reinhardtii* (strain J138)  
C/Species: *Chlamydomonas reinhardtii* subsp. *recombinans*  
C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 23-Mar-2001  
A/Accession: B86498  
R/Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ishiguro, M.; J. Biol. Chem. 276, 2311-2314, 2000  
A/Title: Comparison of whole genome sequences of *Chlamydomonas reinhardtii* J138.  
A/Reference number: A86491; MUID:20330349; PMID:10871362  
A/Accession: B86498  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-207 <STO>  
A/Cross-references: GB:BA000008; NID:9878431; PIDN:BA98268.1; GSPDB:GN00142  
A/Experimental source: strain J138  
C/Genetics:  
A/Gene: *sodM*  
C/Superfamily: superoxide dismutase (Mn)

Query Match 20.0%; Score 6; DB 2; Length 207;  
Best Local Similarity 100.0%; Pred. No. 34;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YSLPEL 7  
|||||  
Db 6 YSLPEL 11

RESULT 14  
B72124

superoxide dismutase (EC 1.15.1.1) (Mn) CP0718 [similarity] - *Chlamydomonas reinhardtii*  
C/Species: *Chlamydomonas reinhardtii* subsp. *recombinans*  
C/Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 26-May-2000  
C/Accession: B72124; B81547  
R/Kalman, S.; Mitchell, W.; Marathe, R.; Lammell, C.; Fan, J.; Olinger, L.; Grimwood, J. Nature Genet. 21, 385-389, 1999  
A/Title: Comparative genomics of *Chlamydomonas reinhardtii* and *C. trachomatis*.  
A/Reference number: A72000; MUID:99206606; PMID:10192388  
A/Accession: B72124  
A/Molecule type: DNA  
A/Residues: 1-207 <ARN>  
A/Cross-references: GB:AE001591; GB:AE001363; NID:94376311; PIDN:AA018210.1; PID:943763  
A/Experimental source: strain CW029  
R/Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, J.; Dodson, R.; Gwinn, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzberg, Nucleic Acids Res. 28, 1397-1406, 2000  
A/Title: Genome sequences of *Chlamydomonas reinhardtii* and *Chlamydomonas trachomatis*.  
A/Reference number: A81500; MUID:20150255; PMID:10684935  
A/Accession: B81547  
A/Molecule type: DNA  
A/Residues: 1-207 <REA>  
A/Cross-references: GB:AE002230; GB:AE002161; NID:97189624; PIDN:AA038524.1; PID:971896  
A/Experimental source: strain AR39, HL cells  
C/Genetics:  
A/Gene: *sodM*; CP0718  
C/Superfamily: superoxide dismutase (Mn)  
C/Keywords: manganese; metalloprotein; oxidoreductase  
F/30,78,166,170/Binding site: manganese (His, His, Asp, His) #status predicted

Query Match 20.0%; Score 6; DB 2; Length 207;  
Best Local Similarity 100.0%; Pred. No. 34;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YSLPEL 7  
|||||  
Db 6 YSLPEL 11

RESULT 15  
AH1190  
3-methyladenine DNA glycosylase homolog lmo0928 [imported] - *Listeria monocytogenes* (str. 3-methyladenine)  
C/Species: *Listeria monocytogenes*  
C/Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 14-Dec-2001  
C/Accession: AH1190  
R/Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baguer, F.; Berche, P.; Bloeker, D.; Jones, L.M.; Karst, U. Science 294, 849-852, 2001  
A/Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Moutouram, A.; Mook, C.; Schlatterer, T.; Simoes, N.; Tietz, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A/Title: Comparative genomics of *Listeria monocytogenes*.  
A/Reference number: AB1077; MUID:21537279; PMID:11679669  
A/Accession: AH1190  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-207 <GLA>  
A/Cross-references: GB:NC\_003210; PIDN:CA099006.1; PID:916410330; GSPDB:GN00177  
A/Experimental source: strain EGD-e  
C/Genetics:  
A/Gene: *lmo0928*  
C/Superfamily: *Bacillus subtilis* DNA-3-methyladenine glycosylase homolog yx1J

Query Match 20.0%; Score 6; DB 2; Length 207;  
Best Local Similarity 100.0%; Pred. No. 34;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 PYISGQ 21  
|||||  
Db 191 PYISGQ 196

RESULT 16  
C83587

hypothetical protein PA0462 [imported] - Pseudomonas aeruginosa (strain PA01)  
C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: G63587  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lim,  
L.; Lory, S.; Olson, M.V.  
N:ature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
A:Reference number: A82950; MUID:20437337; PMID:10984043  
A:Accession: G63587  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-234 <STO>  
A:Cross-references: GB:AE004484; GB:AE004091; NID:G9946320; PIDN:AAG03851.1; GSPDB:GN001  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA0462

Query Match 20.0%; Score 6; DB 2; Length 234;  
Best Local Similarity 100.0%; Pred. No. 39;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LPELUD 8  
Db 95 SLPELD 100

RESULT 17  
E64131  
Formyltetrahydrofolate deformylase (EC 3.5.1.10) - Haemophilus influenzae (strain Rd KW2  
C:Species: Haemophilus influenzae  
C:Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 17-Mar-2000  
C:Accession: E64131; PNO606  
R:Reichmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A  
; Goodyne, J.D.; Scott, J.; Shiley, R.; Liu, L.I.; Glodex, A.; Kelley, J.M.; Weidman, J  
; D.M.; Brandon, R.C.; Fine, L.D.; Fritschman, J.L.; Geoghagen, N.S.M.  
Science 269, 496-512, 1995  
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,  
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
A:Reference number: A64000; MUID:95350630; PMID:7542800  
A:Accession: E64131  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-278 <TIGR>  
A:Cross-references: GB:U52833; GB:L42023; NID:G1574432; PIDN:AAC23236.1; PID:G1574433; T  
R:Maskell, D.  
Gene 129, 155-156, 1993  
A:Title: Cloning and sequencing of the Haemophilus influenzae aroA gene.  
A:Reference number: JN0758; MUID:93328119; PMID:8335255  
A:Accession: PNO606  
A:Molecule type: DNA  
A:Residues: 64-114, 'RNR', 118-137, 'PK', 141-204, 'E', 206-278 <MAS>  
A:Cross-references: GB:L04686; NID:G148863; PIDN:AAA24942.1; PID:G148864  
C:Genetics:  
A:Gene: purN  
C:Function:  
A:Description: catalyzes hydrolysis of 10-formyltetrahydrofolate to tetrahydrofolate and  
A:Pathway: one-carbon metabolism  
A>Note: activated by methionine and inhibited by glycine  
C:Superfamily: phosphoribosylglycinamide formyltransferase; phosphoribosylglycinamide fo  
C:Keywords: hydrolase  
F:85-277/Domain: phosphoribosylglycinamide formyltransferase homology <PRGF>  
F:223/Active site: Asp #status predicted

Query Match 20.0%; Score 6; DB 2; Length 278;  
Best Local Similarity 100.0%; Pred. No. 46;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KYSLPE 6  
Db 66 KYSLPE 71

RESULT 18  
T06212  
glucose and ribitol dehydrogenase homolog - barley  
C:Species: Hordeum vulgare (barley)  
C:Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 11-May-2000  
C:Accession: T06212  
R:Alexander, R.; Alamilli, J.M.; Salami, F.; Bartels, D.  
Planta 192, 519-525, 1994  
A:Title: A novel embryo-specific barley cDNA clone encodes a protein with homologies to  
A:Reference number: Z15538; MUID:94198754; PMID:7764620  
A:Accession: T06212  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-293 <ALE>  
A:Cross-references: EMBL:S7926; NID:G633889; PIDN:AAC60580.1; PID:G633890  
A:Experimental source: cv. Aura, embryo  
C:Genetics:  
A:introns: 34/3  
A:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology  
F:41-228/Domain: short-chain alcohol dehydrogenase homology <SAD>

Query Match 20.0%; Score 6; DB 2; Length 293;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 17 YISGOI 22  
Db 277 YISGOI 282

RESULT 19  
H87192  
Probable isomerase/racemase [imported] - Mycobacterium leprae  
C:Species: Mycobacterium leprae  
C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
C:Accession: H87192  
R:Coile, S.T.; Eigmeyer, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; H  
eam, M.A.; Rutherford, K.M.  
Nature 409, 1007-1011, 2001  
A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; S  
A:Title: Massive gene decay in the leprosy bacillus.  
A:Reference number: A86909; MUID:21128732; PMID:11234002  
A:Accession: H87192  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-334 <STO>  
A:Cross-references: GB:AL450380; NID:G13093906; PIDN:CAC31784.1; GSPDB:GN00147  
C:Genetics:  
A:Gene: ML2268

Query Match 20.0%; Score 6; DB 2; Length 334;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 LPELDY 9  
Db 252 LPELDY 257

RESULT 20  
D86815  
hypothetical protein purM [imported] - Lactococcus lactis subsp. lactis (strain IL1403)  
C:Species: Lactococcus lactis subsp. lactis  
C:Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 03-Aug-2001  
C:Accession: D86815  
R:Botstein, A.; Winkler, P.; Mauger, S.; Jallion, O.; Malarme, K.; Weissenbach, J.; Ehrh  
Genome Res. 11, 731-753, 2001  
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s  
A:Reference number: A86825; MUID:21255186; PMID:11337471  
A:Accession: D86815  
A:Status: preliminary

A;Molecule type: DNA  
A;Residues: 1-338 <STO>  
A;Cross-references: GB:AB005176; PID:g12724523; PIDN:AAK05622.1; GSPDB:GN00146  
A;Experimental source: strain IL1403  
C;Genetics:  
A;Gene: purM  
C;Superfamily: phosphoribosylformylglycinamide cyclo-ligase; phosphoribosylformylglyc

Query Match 20.0%; Score 6; DB 2; Length 338;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SLPELD 8  
DB 208 SLPELD 213

RESULT 21  
E71339  
Inserved hypothetical protein TP0307 - syphilis spirochete  
C;Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 05-Nov-1999  
C;Accession: E71339  
R;Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin  
rson, J.; Khailak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Uettermack, T.; MCD  
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.  
Science 281, 375-388, 1998  
A;Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.  
A;Reference number: A71250; MUID:98332770; PMID:9665876  
A;Accession: E71339  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-342 <COL>  
A;Cross-references: GB:AB001211; GB:AB000520; NID:g3322582; PIDN:AA65296.1; PID:g332258  
A;Experimental source: strain Nichols  
C;Genetics:  
A;Gene: TP0307

Query Match 20.0%; Score 6; DB 2; Length 342;  
Best Local Similarity 100.0%; Pred. No. 56;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YSLPEL 7  
DB 187 YSLPEL 192

RESULT 22  
5408  
Hypothetical protein T28C6.5 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Sep-2000  
C;Accession: T25408  
R;Lloyd, C.  
Submitted to the EMBL Data Library, September 1995  
A;Reference number: Z20030  
A;Accession: T25408  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-345 <WIL>  
A;Cross-references: EMBL:Z54238; PIDN:CAA90996.1; GSPDB:GN00022; CESP:T28C6.5  
A;Experimental source: clone T28C6  
C;Genetics:  
A;Gene: CESP:T28C6.5  
A;Map position: 4  
A;Intons: 44/2; 156/1; 201/1; 238/3; 286/2  
C;Superfamily: Caenorhabditis elegans hypothetical protein BR10.1

Query Match 20.0%; Score 6; DB 2; Length 345;  
Best Local Similarity 100.0%; Pred. No. 56;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LDYEPS 12

DB 265 LDYEPS 270

RESULT 23  
DB4329  
heme biosynthesis protein [imported] - Halobacterium sp. NRC-1  
C;Species: Halobacterium sp. NRC-1  
C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C;Accession: DB4329  
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky,  
J.; Lettner, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl  
ung, K.H.; Alam, M.; Freilae, T.  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; I  
A;Title: Genome sequence of Halobacterium species NRC-1.  
A;Reference number: A84160; MUID:20504483; PMID:11016950  
A;Accession: DB4329  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-352 <STO>  
A;Cross-references: GB:AB004437; NID:g10581232; PIDN:AA620000.1; GSPDB:GN00138  
C;Genetics:  
A;Gene: ntrH

Query Match 20.0%; Score 6; DB 2; Length 352;  
Best Local Similarity 100.0%; Pred. No. 57;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 EFSATE 15  
DB 43 EFSATE 48

RESULT 24  
G83003  
A / G specific adenine glycosylase PA5147 [imported] - Pseudomonas aeruginosa (strain F  
C;Species: Pseudomonas aeruginosa  
C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C;Accession: G83003  
R;Stover, C.K.; Pham, X.Q.; Errin, A.L.; Micooguchi, S.D.; Warrenner, P.; Hickey, M.J.; E  
adman, S.; Yuan, Y.; Brody, L.V.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lathig, K.; Lin  
.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path  
A;Reference number: A82950; MUID:20437337; PMID:10984043  
A;Accession: G83003  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-355 <STO>  
A;Cross-references: GB:AB004927; GB:AB004091; NID:g9951437; PIDN:AA608532.1; GSPDB:GN00  
A;Experimental source: strain PA01  
C;Genetics:  
A;Gene: mutY; PA5147  
C;Superfamily: A/G-specific adenine glycosylase

Query Match 20.0%; Score 6; DB 2; Length 355;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SLPELD 8  
DB 261 SLPELD 266

RESULT 25  
T08277  
carotenoid biosynthesis protein homolog H0660 - Halobacterium sp. (strain NRC-1) plasmid  
N;Alternate names: hypothetical protein H1696  
C;Species: Halobacterium sp.  
A;Variety: strain NRC-1  
C;Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 31-Jan-2000  
C;Accession: T08277; T08366

R;Ng, W.V.; Clinfo, S.A.; Smith, T.M.; Bumgarner, R.E.; Baekin, D.; Faust, J.; Hall, B.;  
Genome Res. 8, 1131-1141, 1998  
A:Title: Snapshot of a large dynamic replicon in a halophilic Archaeon: megaplasmid or m  
A:Reference number: Z16408; MUID:99063795; PMID:9847077  
A:Accession: T08277  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-379 <NGM>  
A:Cross-references: EMBL:AF016485; NID:g2822278; PID:g2822338; HALOSP:H0660  
A:Experimental source: strain NRC-1  
A:Genetics: COP1  
A:Accession: T08366  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-379 <DAS>  
A:Cross-references: EMBL:AF016485; NID:g2822278; PID:g2822427; HALOSP:HI696  
A:Experimental source: strain NRC-1  
A:Genetics: COP2  
C:Genetics: <COP1>  
A:Gene: HALOSP:H0660  
A:Genetics: <COP2>  
A:Genome: plasmid pNRC100  
C:Superfamily: carotenoid biosynthesis protein homolog

Query Match 20.0%; Score 6; DB 2; Length 379;  
Best Local Similarity 100.0%; Pred. No. 62;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 LPELDY 9  
|||||  
Db 61 LPELDY 66

RESULT 26  
G84650  
probable beta-glucosidase [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 17-May-2002  
C:Accession: G84650  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffett, K.S.; Cronin, L.A.; Shen, M.; Vankken, S.E.; Unayam, L.; Tallon, L.;  
Euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487; PMID:10617157  
A:Accession: G84650  
A:Status: preliminary  
A:Residues: 1-384 <STO>  
A:Genetics: DNA  
C:Genetics: DNA  
A:Cross-references: GB:A802093; NID:g4874302; PID:AA031364.1; GSPDB:GN00139  
A:Gene: At2g25630  
A:Map position: 2  
C:Superfamily: Agrobacterium beta-glucosidase

Query Match 20.0%; Score 6; DB 2; Length 384;  
Best Local Similarity 100.0%; Pred. No. 62;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 ATEPYI 18  
|||||  
Db 232 ATEPYI 237

RESULT 27  
F70591  
probable kefB protein - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
C:Accession: F70591  
R:Cole, S.T.; Broech, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.

; Connor, R.; Davies, R.; Devlin, K.; Fellwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.  
Rajadram, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A:Reference number: A70500; MUID:98295987; PMID:9654250  
A:Accession: F70591  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-385 <COL>  
A:Cross-references: GB:Z95121; GB:AL123456; NID:g3261742; PID:CA08337.1; PID:g2072703  
A:Experimental source: strain H37Kv  
C:Genetics:  
A:Gene: kefB  
C:Superfamily: Aquifex aeolicus Na+/H+-exchanging protein napA1

Query Match 20.0%; Score 6; DB 2; Length 385;  
Best Local Similarity 100.0%; Pred. No. 63;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 EFSATE 15  
|||||  
Db 76 EFSATE 81

RESULT 28  
G87006  
probable transmembrane transport protein ML0782 [imported] - Mycobacterium leprae  
C:Species: Mycobacterium leprae  
C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 17-May-2002  
C:Accession: G87006  
R:Cole, S.T.; Eigmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; H  
R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Fellwell, T.; Fraser, A.; Hamlin, N.; Holroyd  
eam, M.A.; Rutherford, K.M.  
Nature 409, 1007-1011, 2001  
A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; S  
A:Title: Massive gene decay in the leprosy bacillus.  
A:Reference number: A86909; MUID:21128732; PMID:11234002  
A:Accession: G87006  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-385 <STO>  
A:Cross-references: GB:AL450380; NID:g13092891; PID:CA030291.1; GSPDB:GN00147  
C:Genetics:  
A:Gene: ML0782  
C:Superfamily: Aquifex aeolicus Na+/H+-exchanging protein napA1

Query Match 20.0%; Score 6; DB 2; Length 385;  
Best Local Similarity 100.0%; Pred. No. 63;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 EFSATE 15  
|||||  
Db 76 EFSATE 81

RESULT 29  
A32812  
repa protein - Agrobacterium tumefaciens plasmid pTiBES3  
C:Species: Agrobacterium tumefaciens  
C:Date: 20-Oct-1989 #sequence\_revision 28-Aug-1992 #text\_change 08-Oct-1999  
C:Accession: A32812  
R:Tabata, S.; Hooykaas, P.J.J.; Oka, A.  
J. Bacteriol. 171, 1665-1672, 1989  
A:Title: Sequence determination and characterization of the replicator region in the tu  
A:Reference number: A32812; MUID:89155477; PMID:2537824  
A:Accession: A32812  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-405 <TAB>  
A:Cross-references: GB:M24529; NID:g154803; PID:AAA27402.1; PID:g154804  
C:Genetics:  
A:Genome: plasmid



## C:Superfamily: sopA protein

Query Match 20.0%; Score 6; DB 2; Length 405;  
Best Local Similarity 100.0%; Pred. No. 66;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 QOINBI 25  
Db 90 QOINBI 95

## RESULT 30

AI3332  
replication protein A [imported] - Agrobacterium tumefaciens (strain C58, Dupont) plasmid  
C/Species: Agrobacterium tumefaciens  
C/Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 01-Feb-2002  
C/Accession: AI3332

R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monke, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell  
Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001

A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
ster, E.W.

A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A;Reference number: AB2577; PMID:11743193

A;Accession: AI3332

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-405 <KUR>

A;Cross-references: GB:AE008690; PIDN:AAI46279.1; PID:917744061; GSPDB:GN00189

A;Experimental source: strain C58 (Dupont)

C;Genetics:

A;Gene: repA

A;Superfamily: sopA protein

Query Match 20.0%; Score 6; DB 2; Length 405;  
Best Local Similarity 100.0%; Pred. No. 66;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 QOINBI 25  
Db 90 QOINBI 95

## RESULT 31

D96939

theory protein, containing EAL-domain [imported] - Clostridium acetobutylicum

C/Species: Clostridium acetobutylicum

C/Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 14-Sep-2001

C/Accession: D96939

R;Nolling, J.; Bretton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,  
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001

A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium C1C

A;Reference number: A96900; MUID:21359325; PMID:21359325

A;Accession: D96939

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-412 <KUR>

A;Cross-references: GB:AE001437; PIDN:AAK78303.1; PID:915023166; GSPDB:GN00168

A;Experimental source: Clostridium acetobutylicum ATCC824

C;Genetics:

A;Gene: CAC0322

Query Match 20.0%; Score 6; DB 2; Length 412;  
Best Local Similarity 100.0%; Pred. No. 67;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 TEPIYS 19  
Db 372 TEPIYS 377

## RESULT 32

C64080

hemY protein homolog - Haemophilus influenzae (strain Rd KW20)

C/Species: Haemophilus influenzae

C/Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 08-Oct-1995

C/Accession: C64080

R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kiehlavage,  
Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman,  
D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.

Science 269, 496-512, 1995

A;Authors: Gnehm, C.U.; McDonald, L.A.; Small, K.V.; Frazer, C.M.; Smith, H.O.; Venter,  
A;Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.

A;Reference number: A64000; MUID:95350630; PMID:7542800

A;Accession: C64080

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-428 <TIGR>

A;Cross-references: GB:U32742; GB:I42023; NID:91573593; PIDN:AAK22260.1; PID:91573594;

Query Match 20.0%; Score 6; DB 2; Length 428;  
Best Local Similarity 100.0%; Pred. No. 69;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 PYISGQ 21  
Db 19 PYISGQ 24

## RESULT 33

T50362

cdc7-like protein kinase [imported] - fission yeast (Schizosaccharomyces pombe)

C/Species: Schizosaccharomyces pombe

C/Date: 09-Jun-2000 #sequence\_revision 09-Jun-2000 #text\_change 09-Jun-2000

C/Accession: T50362

R;Seeger, K.; Harris, D.; Lyne, M.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, February 2000

A;Reference number: Z25064

A;Accession: T50362

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-429 <SEE>

A;Cross-references: EMBL:AL157918; PIDN:CA876054.1; GSPDB:GN00067; SPDB:SPBC21C3.18

A;Experimental source: strain 972h(-); cosmid c21C3

C;Genetics:

A;Gene: SPDB:SPBC21C3.18

A;Map position: 2

A;Intons: 332/1; 388/1

Query Match 20.0%; Score 6; DB 2; Length 429;  
Best Local Similarity 100.0%; Pred. No. 70;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KYSLPE 6  
Db 153 KYSLPE 158

## RESULT 34

AH2930

oxido-reductase Atu3046 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)

C/Species: Agrobacterium tumefaciens

C/Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 11-Jan-2002

C/Accession: AH2930

R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monke, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo,  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell  
Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
ster, E.W.

A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A;Reference number: AB2577; PMID:11743193

A;Accession: AH2930

A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-441 <KUR>  
A:Cross-references: GB:AE008689; PIDN:AA143862.1; PID:GL7741407; GSPDB:GN00187  
A:Experimental source: Strain C58 (Dupont)  
C:Genetics:  
A:Gene: Atcu3046  
A:Map position: linear chromosome

Query Match 20.0%; Score 6; DB 2; Length 441;  
Best Local Similarity 100.0%; Pred. No. 71;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 YSLPEL 7  
|||||  
Db 83 YSLPEL 88

## RESULT 35

1101  
Species: Haemophilus ducreyi (Fragment)

C>Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 02-Mar-2001  
C/Accession: J11101  
R/Ward, C.K.; Lumbley, S.R.; Latimer, J.L.; Cope, L.D.; Hansen, E.J.  
J. Bacteriol. 180, 6013-6022, 1998  
A:Title: Haemophilus ducreyi secretes a filamentous hemagglutinin-like protein.  
A:Reference number: Z20984; MUID:99030326; PMID:9811662  
A:Accession: J11101  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-443 <MAR>  
A:Cross-references: EMBL:AF057695; NID:g3929017; PID:g3929019; PIDN:AACT9758.1  
C:Genetics:  
A:Gene: pmm  
C:Superfamily: Mycoplasma pitum phosphomannomutase

Query Match 20.0%; Score 6; DB 2; Length 443;  
Best Local Similarity 100.0%; Pred. No. 72;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 17 YISGQI 22  
|||||  
Db 331 YISGQI 336

## RESULT 36

98351  
Species: Agrobacterium tumefaciens (Strain C58, Cere

C>Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 11-Jan-2002  
C/Accession: P98351.  
R/Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quirillo, B.; Goldman, A.; Liu, F.; Mollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001  
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens.  
A:Reference number: A97359; PMID:11743194  
A:Accession: P98351  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-446 <KUR>  
A:Cross-references: GB:AE007870; PIDN:AAK90336.1; PID:GL5160373; GSPDB:GN00170  
C:Genetics:  
A:Gene: AGR\_L\_3517  
A:Map position: linear chromosome

Query Match 20.0%; Score 6; DB 2; Length 446;  
Best Local Similarity 100.0%; Pred. No. 72;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 YSLPEL 7  
|||||  
Db 88 YSLPEL 93

## RESULT 37

530839  
UTR2 protein - yeast (Saccharomyces cerevisiae)

N/Alternate names: protein YEL040w  
C/Species: Saccharomyces cerevisiae  
C>Date: 28-May-1993 #sequence\_revision 28-May-1993 #text\_change 06-Feb-1998  
A:Accession: S30839; S50504; S38545  
R/Mulligan, J.T.; Dietrich, F.S.; Hennessey, K.M.; Seh1, P.; Komp, C.; Wei, Y.; Taylor, submitted to the EMBL Data Library, February 1993  
A:Reference number: S30812  
A:Accession: S30839

A:Molecule type: DNA  
A:Residues: 1-467 <MUL>  
A:Cross-references: GB:U18779; EMBL:U10830; NID:g603625; PID:g603639  
R/Dietrich, F.S. EMBL Data Library, December 1994  
submitted to the EMBL Data Library, December 1994  
A:Description: The sequence of S. cerevisiae cosmid 8199, 8334, and 9871.  
A:Reference number: S50491  
A:Accession: S50504

A:Molecule type: DNA  
A:Residues: 1-467 <DIE>  
A:Cross-references: EMBL:U18779; NID:g603625; PID:g603639; MIPS:YEL040w  
R/Melnick, L.; Sherman, F.  
J. Mol. Biol. 233, 372-388, 1993  
A:Title: The gene clusters ARC and COR on chromosomes 5 and 10, respectively, of Saccha  
A:Reference number: S38543; MUID:94016558; PMID:8411151  
A:Accession: S38545

A:Molecule type: DNA  
A:Residues: 121-129, 'V', 131-290, 'R', 292-353, 'C', 355-467 <MEL>  
A:Cross-references: EMBL:S66130; NID:g430829; PID:g430830  
C:Genetics:  
A:Gene: SGD:UTR2  
A:Cross-references: SGD:S0000766; MIPS:YEL040w  
A:Map position: 5L

Query Match 20.0%; Score 6; DB 2; Length 467;  
Best Local Similarity 100.0%; Pred. No. 75;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 ELDYEF 11  
|||||  
Db 166 ELDYEF 171

## RESULT 38

E97012  
probable non-processive endoglucanase family 5, secreted, Cella homolog secreted, docker

C/Species: Clostridium acetobutylicum  
C>Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 14-Sep-2001  
C/Accession: E97012  
R/Nolling, U.; Berton, G.; Omeichenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee  
J. Bacteriol. 183, 4823-4838, 2001  
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium C1  
A:Reference number: A96900; MUID:21359325; PMID:21359325  
A:Accession: E97012  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-482 <KUR>  
A:Cross-references: GB:AE001437; PIDN:AAK78888.1; PID:GL5023812; GSPDB:GN00168  
A:Experimental source: Clostridium acetobutylicum ATCC824  
C:Genetics:  
A:Gene: CAC0912

Query Match 20.0%; Score 6; DB 2; Length 482;  
Best Local Similarity 100.0%; Pred. No. 78;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 17 YISGQI 22  
|||||  
Db 445 YISGQI 450

Job time : 37.2936 secs

## RESULT 39

F84001  
glycine dehydrogenase subunit 2 BH2814 [imported] - Bacillus halodurans (strain C-125)  
C/Species: Bacillus halodurans  
C/Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001  
C/Accession: F84001  
R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira  
Nucleic Acids Res. 28, 4317-4331, 2000  
A/Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
A/Reference number: A83650; MUID:20512582; PMID:11058132  
A/Accession: F84001  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-488 <STO>  
A/Cross-references: GB:AP001516; GB:BA000004; NID:G10175192; PIDN:BA06553.1; GSPDB:GN00  
A/Experimental source: strain C-125  
C/Genetics:  
Gene: BH2814

## Query Match

20.0%; Score 6; DB 2; Length 488;  
Best Local Similarity 100.0%; Pred. No. 79;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SLPELD 8

Db 21 SLPELD 26

## RESULT 40

T02400  
Probable beta-glucosidase [imported] - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 05-Mar-1999 #sequence\_revision 05-Mar-1999 #text\_change 16-Feb-2001  
C/Accession: T02400; F84878  
R/Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul  
submitted to the EMBL Data Library, May 1998  
A/Description: Arabidopsis thaliana chromosome II BAC F411 genomic sequence.  
A/Reference number: Z14667  
A/Accession: T02400  
A/Status: translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-506 <ROU>

A/Cross-references: EMBL:AC004521; NID:G3128166; PIDN:AC16091.1; PID:G3128187

A/Experimental source: cultivar Columbia  
R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shee, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.;  
; D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
ature 402, 761-768, 1999  
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A/Reference number: A84420; MUID:20083487; PMID:10617197

A/Accession: F84878

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-506 <STO>

A/Cross-references: GB:AE002093; NID:G3128187; PIDN:AC16091.1; GSPDB:GN00139

C/Genetics:  
A/Gene: F411.26; At2g44450

A/Map position: 2

A/Intons: 50/3; 74/1; 93/3; 119/1; 145/1; 174/2; 259/3; 370/1; 380/3; 415/1; 447/2  
C/Superfamily: Agrobacterium beta-glucosidase

## Query Match

20.0%; Score 6; DB 2; Length 506;  
Best Local Similarity 100.0%; Pred. No. 82;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 ATEPYI 18

Db 233 ATEPYI 238

Search completed: April 9, 2003, 14:17:53



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 8, 2003, 02:14:28 ; Search time 9.3578 Seconds  
(without alignments)  
983.170 Million cell updates/sec

Title: US-09-987-190-2  
Perfect score: 145  
Sequence: 1 KXSLPELDYFSATPEYISQINIEIXYTX 30

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Command line parameters:

-MODE=frame+ p2n.model -DEV=x1h  
-O=/cgn2\_1/USPTO\_spool/US09987190/runat\_02042003\_092633\_19253/app\_query.fasta\_1.526  
-DB=Issued\_Patents\_NA -OPMT=fastac -SOFTX=tni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40  
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09987190 @CGN 1.1 47 @runat\_02042003\_092633\_19253 -NCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT\_LONGLOG -DEV TIMEOUT=120  
-MAEN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

## Database :

- 1: Issued Patents\_NA.\*
- 2: /cgn2\_6/prodata/1/ina/5A COMB.seq.\*
- 3: /cgn2\_6/prodata/1/ina/5B COMB.seq.\*
- 4: /cgn2\_6/prodata/1/ina/6A COMB.seq.\*
- 5: /cgn2\_6/prodata/1/ina/6B COMB.seq.\*
- 6: /cgn2\_6/prodata/1/ina/backfile1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	79	54.5	812	4	US-09-091-097-7
2	74	51.0	594	4	US-09-075-019-1
3	74	51.0	600	6	5240847-8
4	74	51.0	600	6	5240847-9
5	74	51.0	600	6	5240847-16
6	74	51.0	600	6	5240847-17
7	74	51.0	681	4	US-09-075-019-6
8	74	51.0	813	2	US-08-927-230A-1
9	74	51.0	813	3	US-09-151-052-1
10	74	51.0	969	2	US-08-365-486A-27
11	74	51.0	969	4	US-08-880-342-27
12	74	51.0	976	4	US-09-126-109-1

13	74	51.0	977	2	US-08-023-980B-2
14	74	51.0	977	2	US-08-486-953A-2
15	74	51.0	987	6	5240847-19
16	74	51.0	1032	6	5240847-25
17	74	51.0	3789	4	US-09-075-019-8
18	74	49.0	721	4	US-09-262-856A-8
19	67	46.2	99	6	5240847-7
20	62	42.8	728	4	US-09-091-097-5
21	58	40.0	719	4	US-09-411-578-39
22	57	39.3	606	4	US-09-134-001C-747
23	53	36.6	780	1	US-08-445-909A-28
24	53	36.6	1294	1	US-08-445-909A-16
25	50.5	34.8	747	2	US-08-894-772-4
26	50.5	34.8	747	2	US-09-207-844-4
27	50	34.5	2904	4	US-09-605-785-703
28	50	34.5	3410	4	US-09-020-956-110
29	50	34.5	3410	4	US-09-030-607-110
30	50	34.5	3410	4	US-09-605-785-110
31	50	34.5	3410	4	US-09-439-313-110
32	50	34.5	3410	4	US-09-352-616A-110
33	50	34.5	3410	4	US-09-602-877A-100
34	50	34.5	3410	4	US-09-232-149A-110
35	50	34.5	4034	4	US-09-605-785-704
36	50	34.5	4137	4	US-09-221-017B-329
37	50	34.5	4894	4	US-09-605-785-702
38	50	34.5	6976	4	US-09-605-785-705
39	49	33.8	99500	4	US-09-798-096-10
40	48.5	33.4	2943	4	US-09-735-935-3
41	48	33.1	2943	1	US-08-042-747B-7
42	48	33.1	3572	2	US-08-713-815A-2
43	48	33.1	4080	2	US-08-446-345-35
44	48	33.1	8931	3	US-09-028-934-28
45	47.5	32.8	1769	4	US-09-257-584-6

## ALIGNMENTS

RESULT 1  
US-09-091-097-7  
; Sequence 7, Application US/09091097  
; Patent No. 6432407  
; GENERAL INFORMATION:  
; APPLICANT: TAKESAKO, KAZUTOH  
; APPLICANT: OKADO, TAKASHI  
; APPLICANT: YAGIHARA, TOMOKO  
; APPLICANT: KURODA, MASANOBU  
; APPLICANT: ONISHI, YOSHIMI  
; APPLICANT: KATO, IKUNOSHIN  
; APPLICANT: AKIYAMA, KAZUO  
; APPLICANT: YASUEDA, HIROSHI  
; APPLICANT: YAMAGUCHI, HIDEYO  
; TITLE OF INVENTION: ANTIGENIC PROTEIN ORIGINATING IN  
; NUMBER OF SEQUENCES: 58  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH, LLP  
; STREET: PO BOX 747  
; CITY: FALLS CHURCH  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22040-0747  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/091, 097  
; FILING DATE:  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WEINER, MARC S.  
; REGISTRATION NUMBER: 32,181

REFERENCE/DOCKET NUMBER: 1422-0346P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-205-8000  
TELEFAX: 703-205-8050  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 812 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 2..673  
US-09-091-097-7  
Alignment Scores:  
Pred. No.: 1.05e-05 Length: 812  
Score: 79.00 Matches: 15  
Percent Similarity: 70.37% Conservative: 8  
Best Local Similarity: 55.56% Mismatches: 8  
Query Match: 54.48% Indels: 0  
Gaps: 0  
US-09-987-190-2 (1-30) x US-09-091-097-7 (1-812)  
Oy 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20  
Db 74 AAGTACAGCGCTGCCCGCTGCGTACGACTAGCGCGCTGACCGCGCATCTCGGAC 133  
Oy 21 GlnIleAsnGluIle\*\*Tyr 27  
Db 134 GAGATCATGAGAGACGACTAC 154  
RESULT 2  
US-09-075-019-1  
Sequence 1, Application US/09075019  
Patent No. 6190658  
GENERAL INFORMATION:  
APPLICANT: UTC IR459  
TITLE OF INVENTION: A GENETICALLY MODIFIED MANGANESE  
NUMBER OF SEQUENCES: 12  
TITLE OF INVENTION: SUPEROXIDE DISMUTASE FOR TREATING OXIDATIVE DAMAGE  
CORRESPONDENCE ADDRESSES: 12  
ADDRESSEE: Sheridan Rose P. C.  
STREET: 1700 Lincoln St., Suite 3500  
CITY: Denver  
STATE: CO  
COUNTRY: USA  
ZIP: 80203  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/075,019  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Kovarik, Joseph E.  
REGISTRATION NUMBER: 33,005  
REFERENCE/DOCKET NUMBER: 2848-22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 863-9700  
FAX: (303) 863-0223  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 594 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA

FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..594  
US-09-075-019-1  
Alignment Scores:  
Pred. No.: 6.04e-05 Length: 594  
Score: 74.00 Matches: 12  
Percent Similarity: 78.57% Conservative: 10  
Best Local Similarity: 42.86% Mismatches: 6  
Query Match: 51.03% Indels: 0  
Gaps: 0  
US-09-987-190-2 (1-30) x US-09-075-019-1 (1-594)  
Oy 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20  
Db 1 AAGCAGCGCTGCCCGACTGCGCTACGACTAGCGCGCTGGAACCTCACATCAACGCG 60  
Oy 21 GlnIleAsnGluIle\*\*TyrThr 28  
Db 61 CAGATCATGACGCTGACACACAGC 84  
RESULT 3  
5240847-8  
Patent No. 5240847  
APPLICANT: HECKL, KONRAD; SPEVAK, WALTER; OSTERMANN, ELINBORG;  
ZOEHEL, ANDREAS; KRYSTEK, EDELTRAUD; MAURER-FOGY, INGRID;  
WITCHE-CASPARON, MARIA J.; STRATOWA, CHRISTIAN; HAUPTMANN, RUDOLF  
TITLE OF INVENTION: HUMAN MANGANESE SUPEROXIDE DISMUTASE  
(HNN-SOD)  
NUMBER OF SEQUENCES: 34  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/167,261  
FILING DATE: 11-MAR-1988  
SEQ ID NO: 8  
LENGTH: 600  
5240847-8  
Alignment Scores:  
Pred. No.: 6.12e-05 Length: 600  
Score: 74.00 Matches: 12  
Percent Similarity: 78.57% Conservative: 10  
Best Local Similarity: 42.86% Mismatches: 6  
Query Match: 51.03% Indels: 0  
Gaps: 0  
US-09-987-190-2 (1-30) x 5240847-8 (1-600)  
Oy 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20  
Db 4 AAGCAGCTTTGCCAGACTTGCCATAGCAGTACGCGTCTAGAACACACATCAATGCT 63  
Oy 21 GlnIleAsnGluIle\*\*TyrThr 28  
Db 64 CAATCATGCAATGACACACTCT 87  
RESULT 4  
5240847-9  
Patent No. 5240847  
APPLICANT: HECKL, KONRAD; SPEVAK, WALTER; OSTERMANN, ELINBORG;  
ZOEHEL, ANDREAS; KRYSTEK, EDELTRAUD; MAURER-FOGY, INGRID;  
WITCHE-CASPARON, MARIA J.; STRATOWA, CHRISTIAN; HAUPTMANN, RUDOLF  
TITLE OF INVENTION: HUMAN MANGANESE SUPEROXIDE DISMUTASE  
(HNN-SOD)  
NUMBER OF SEQUENCES: 34  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/167,261  
FILING DATE: 11-MAR-1988  
SEQ ID NO: 9  
LENGTH: 600  
5240847-9



Db 67 CAGATCATGACGTGACACACAGC 90

RESULT 8

US-08-927-230A-1

Sequence 1, Application US/08927230A

Patent No. 5985633

GENERAL INFORMATION:

APPLICANT: Nick et al.

TITLE OF INVENTION: Human Manganese Superoxide Dismutase Mutants

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD, LLP

STREET: 28 State Street, Floor 24

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109-1875

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/927,230A

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Remillard, Jane

REGISTRATION NUMBER: 38,872

REFERENCE/DOCKET NUMBER: UFI-001

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)227-7400

TELEFAX: (617)742-4214

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 813 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 43..708

FEATURE:

NAME/KEY: mat\_peptide

LOCATION: 115..708

US-08-927-230A-1

Alignment Scores:

Pred. No.: 9.09e-05

Score: 74.00

Percent Similarity: 78.57%

Best Local Similarity: 42.86%

Query Match: 51.03%

Length: 813

Matches: 12

Conservative: 10

Mismatches: 6

Indels: 0

Gaps: 0

US-09-987-190-2 (1-30) x US-08-927-230A-1 (1-813)

Qy 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20

Db 115 AAGCACACGCTCCCGACCTGCGCTACGACTACGCGCGCCCTGGAACCTCATCAACGCG 174

Qy 21 GlnIleAsnGluIle\*\*TyrThr 28

Db 175 CAGATCATGACGTGACACACAGC 198

RESULT 9

US-09-151-052-1

Sequence 1, Application US/09151052

Patent No. 6107070

GENERAL INFORMATION:

APPLICANT: Nick et al.

TITLE OF INVENTION: Human Manganese Superoxide Dismutase Mutants

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD

STREET: 28 State Street, Floor 24

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109-1875

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/151,052

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/927,230

FILING DATE: September 10, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Remillard, Jane

REGISTRATION NUMBER: 38,872

REFERENCE/DOCKET NUMBER: UFI-001

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)227-7400

TELEFAX: (617)742-4214

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 813 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 43..708

FEATURE:

NAME/KEY: mat\_peptide

LOCATION: 115..708

US-09-151-052-1

Alignment Scores:

Pred. No.: 9.09e-05

Score: 74.00

Percent Similarity: 78.57%

Best Local Similarity: 42.86%

Query Match: 51.03%

Length: 813

Matches: 12

Conservative: 10

Mismatches: 6

Indels: 0

Gaps: 0

US-09-987-190-2 (1-30) x US-09-151-052-1 (1-813)

Qy 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20

Db 115 AAGCACACGCTCCCGACCTGCGCTACGACTACGCGCGCCCTGGAACCTCATCAACGCG 174

Qy 21 GlnIleAsnGluIle\*\*TyrThr 28

Db 175 CAGATCATGACGTGACACACAGC 198

RESULT 10

US-08-365-486A-27

Sequence 27, Application US/08365486A

Patent No. 5834306

GENERAL INFORMATION:

APPLICANT: Webster, Keith A.

TITLE OF INVENTION: Tissue Specific Hypoxia Regulated

NUMBER OF SEQUENCES: 31



CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dehlinger & Associates  
STREET: 350 Cambridge Avenue, Suite 250  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/365,486A  
FILING DATE: 23-DEC-1994  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Sholtz, Charles K.  
REGISTRATION NUMBER: 38,615  
REFERENCE/DOCKET NUMBER: 8255-0018  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 969 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: unknown  
MOLECULE TYPE: cDNA to mRNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: human manganese superoxide dismutase  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 61..729  
US-08-365-486A-27  
Alignment Scores:  
Pred. No.: 0.000114 Length: 969  
Score: 74.00 Matches: 12  
Percent Similarity: 78.57% Conservative: 10  
Best Local Similarity: 42.86% Mismatches: 6  
Query Match: 51.03% Indels: 0  
Gaps: 0  
US-09-987-190-2 (1-30) x US-08-365-486A-27 (1-969)  
QY 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20  
Db 133 AMGCACGACCTCCCGACCTCCCTACGACTACGCGCCCTGGAACCTCACAATCAACGCG 192  
QY 21 GlnIleAsnGluIle\*\*\*TyrThr 28  
Db 193 CAGATCATGCAGCTGCACCAACAGC 216  
RESULT 11  
US-08-880-342-27  
Sequence 27, Application US/08880342  
Patent No. 6218179  
GENERAL INFORMATION:  
APPLICANT: Webster, Keith A.  
APPLICANT: Bishopric, Nanette H.  
APPLICANT: Murphy, Brian  
APPLICANT: Laderoute, Keith R.  
APPLICANT: Green, Christopher J.  
TITLE OF INVENTION: Tissue Specific Hypoxia Regulated  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Dehlinger & Associates

STREET: 350 Cambridge Avenue, Suite 250  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/880,342  
FILING DATE: 23-JUN-1997  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/IB95/00996  
FILING DATE: 13-NOV-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/365,486  
FILING DATE: 23-DEC-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Sholtz, Charles K.  
REGISTRATION NUMBER: 38,615  
REFERENCE/DOCKET NUMBER: 8255-0018.30  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 969 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: unknown  
MOLECULE TYPE: cDNA to mRNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: human manganese superoxide dismutase  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 61..729  
US-08-880-342-27  
Alignment Scores:  
Pred. No.: 0.000114 Length: 969  
Score: 74.00 Matches: 12  
Percent Similarity: 78.57% Conservative: 10  
Best Local Similarity: 42.86% Mismatches: 6  
Query Match: 51.03% Indels: 0  
Gaps: 0  
US-09-987-190-2 (1-30) x US-08-880-342-27 (1-969)  
QY 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20  
Db 133 AMGCACGACCTCCCGACCTCCCTACGACTACGCGCCCTGGAACCTCACAATCAACGCG 192  
QY 21 GlnIleAsnGluIle\*\*\*TyrThr 28  
Db 193 CAGATCATGCAGCTGCACCAACAGC 216  
RESULT 12  
US-09-126-109-1  
Sequence 1, Application US/09126109  
Patent No. 6171856  
GENERAL INFORMATION:  
APPLICANT: Thigpen, Anice  
APPLICANT: Hommeier, Hans-Ewald  
APPLICANT: Newgard, Christopher B.  
APPLICANT: Unger, Roger H.  
APPLICANT: Shimabukuro, Michio  
APPLICANT: Chen, Guaxun

APPLICANT: Rhodes, Christopher J.  
APPLICANT: Hugl, Sharon R.  
APPLICANT: Cousin, Sharon  
TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING  
TO NO-MEDIATED CYTOTOXICITY  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/126,109  
FILING DATE: 30-JUL-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/055,092  
FILING DATE: 30-JUL-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US Unknown  
FILING DATE: 03-MAR-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: McMillian, Nabehla R.  
REGISTRATION NUMBER: P-43,363  
REFERENCE/DOCKET NUMBER: UTSD:560  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (512) 474-7577  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 976 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-09-126-109-1

Alignment Scores:  
Pred. No.: 0.000115 Length: 976  
Score: 74.00 Matches: 12  
Percent Similarity: 78.57% Conservative: 10  
Best Local Similarity: 42.86% Mismatches: 6  
Query Match: 51.03% Indels: 0  
Gaps: 0

US-09-987-190-2 (1-30) x US-09-126-109-1 (1-976)

Qy 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20  
Db 167 AAGCAGACGCTCCCGACTGCTCCCTACGACTACGGCGCCCTGGAACCTCATCATCAACGCG 226

Qy 21 GlnIleAngIuIle\*\*\*TyrThr 28  
Db 227 CAGATCATGACAGTCGACACACAGC 250

RESULT 13  
US-08-023-980B-2  
Sequence 2, Application US/08023980B  
Patent No. 5843641  
GENERAL INFORMATION:  
APPLICANT: Brown, Robert  
APPLICANT: Horvitz, H. Robert  
APPLICANT: Rosen, Daniel R.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS,  
TREATMENT AND PREVENTION OF DISEASES OF CELL DEATH  
NUMBER OF SEQUENCES: 45

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Elbing LLP  
STREET: 585 Commercial Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109-1024  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/023,980B  
FILING DATE: 26-FEB-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 00786/177001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/723-4123  
TELEFAX: 617/723-8962  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 977 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-08-023-980B-2

Alignment Scores:  
Pred. No.: 0.000116 Length: 977  
Score: 74.00 Matches: 12  
Percent Similarity: 78.57% Conservative: 10  
Best Local Similarity: 42.86% Mismatches: 6  
Query Match: 51.03% Indels: 0  
Gaps: 0

US-09-987-190-2 (1-30) x US-08-023-980B-2 (1-977)

Qy 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20  
Db 168 AAGCAGACGCTCCCGACTGCTCCCTACGACTACGGCGCCCTGGAACCTCATCATCAACGCG 227

Qy 21 GlnIleAngIuIle\*\*\*TyrThr 28  
Db 228 CAGATCATGACAGTCGACACACAGC 251

RESULT 14  
US-08-486-953A-2  
Sequence 2, Application US/08486953A  
Patent No. 5849290  
GENERAL INFORMATION:  
APPLICANT: Brown, Robert  
APPLICANT: Horvitz, H. Robert  
APPLICANT: Rosen, Daniel R.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS,  
TREATMENT AND PREVENTION OF DISEASES OF CELL DEATH  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Elbing LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

```
SOFTWARE: FastSeq
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,953A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/204,052
FILING DATE: 28-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/223002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/428-0200
TELEFAX: 617/428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 977 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-486-953A-2

Alignment Scores:
Pred. No.: 0.000116 Length: 977
Score: 74.00 Matches: 12
Percent Similarity: 78.57% Conservative: 10
Best Local Similarity: 42.86% Mismatches: 6
Query Match: 51.03% Indels: 0
Gaps: 0

US-09-987-190-2 (1-30) x US-08-486-953A-2 (1-977)

QY 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20
Db 168 AAGCAGACGCTCCCGACCTGCTACGACTACGCGCCCTGGAACCTCACATCAACCG 227

QY 21 GlnIleAsnGluIle**TyrThr 28
Db 228 CAGATCATGACGCTGCACACACG 251

RESULT 15
5240847-19
PATENT NO. 5240847
APPLICANT: HECKL, KONRAD; SPEVAK, WALTER; OSTERMANN, ELINBORG;
ZEHLE, ANDREAS; KRISTEK, EDELTRAUD; MAURER-FOGY, INGRID;
CHE-CASTANON, MARIA J.; STRATOWA, CHRISTIAN; HAUPTMANN, RUDOLF
TITLE OF INVENTION: HUMAN MANGANESE SUPEROXIDE DISMUTASE
(HMN-SOD)
NUMBER OF SEQUENCES: 34
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/167,261
FILING DATE: 11-MAR-1988
SEQ ID NO: 19
LENGTH: 987
5240847-19

Alignment Scores:
Pred. No.: 0.000117 Length: 987
Score: 74.00 Matches: 12
Percent Similarity: 78.57% Conservative: 10
Best Local Similarity: 42.86% Mismatches: 6
Query Match: 51.03% Indels: 0
Gaps: 0

US-09-987-190-2 (1-30) x 5240847-19 (1-987)

QY 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20
Db 13 AAGACCTCTTGTCCAGACTTGCATACGACTACGGTGTCTTAGAACACACATCATATGCT 72
```

```
QY 21 GlnIleAsnGluIle**TyrThr 28
Db 73 CAATCATGCAATGACACACTCT 96

RESULT 16
5240847-25
PATENT NO. 5240847
APPLICANT: HECKL, KONRAD; SPEVAK, WALTER; OSTERMANN, ELINBORG;
ZEHLE, ANDREAS; KRISTEK, EDELTRAUD; MAURER-FOGY, INGRID;
CHE-CASTANON, MARIA J.; STRATOWA, CHRISTIAN; HAUPTMANN, RUDOLF
TITLE OF INVENTION: HUMAN MANGANESE SUPEROXIDE DISMUTASE
(HMN-SOD)
NUMBER OF SEQUENCES: 34
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/167,261
FILING DATE: 11-MAR-1988
SEQ ID NO: 25
LENGTH: 1032
5240847-25

Alignment Scores:
Pred. No.: 0.000124 Length: 1032
Score: 74.00 Matches: 12
Percent Similarity: 78.57% Conservative: 10
Best Local Similarity: 42.86% Mismatches: 6
Query Match: 51.03% Indels: 0
Gaps: 0

US-09-987-190-2 (1-30) x 5240847-25 (1-1032)

QY 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20
Db 73 AAGCAGACGCTCCCGACCTGCTACGACTACGCGCCCTGGAACCTCACATCAACCG 132

QY 21 GlnIleAsnGluIle**TyrThr 28
Db 133 CAGATCATGACGCTGCACACACG 156

RESULT 17
US-09-075-019-8
Sequence 8, Application US/09075019
PATENT NO. 6190658
GENERAL INFORMATION:
APPLICANT: UTC IR459
TITLE OF INVENTION: A GENETICALLY MODIFIED MANGANESE
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross P.C.
STREET: 1700 Lincoln St., Suite 3500
CITY: Denver
STATE: CO
COUNTRY: USA
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/075,019
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Kovarik, Joseph E.
REGISTRATION NUMBER: 33,005
REFERENCE/DOCKET NUMBER: 2848-22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 3789 base pairs
```

```

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-075-019-8

Alignment Scores:
Pred. No.: 0.000677 Length: 3789
Score: 74.00 Matches: 12
Percent Similarity: 78.57% Conservative: 10
Best Local Similarity: 42.86% Mismatches: 6
Query Match: 51.03% Indels: 0
DB: 4 Gaps: 0

US-09-987-190-2 (1-30) x US-09-075-019-8 (1-3789)

Qy 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20
Db 550 AAGCAGACGCTCCCGACCTGCTACGCGCTGGAACCTCACTCAACGCG 609
21 GlnIleAenGluIle**TyrThr 28
610 CAGATCATGCACTGCACCAACGCG 633

RESULT 18
US-09-262-856A-8
; Sequence 8, Application US/09262856A
; Patent No. 6333164
; GENERAL INFORMATION:
; APPLICANT: TAKESAKO, KAZUTOH
; APPLICANT: MIZUTANI, SHIGETOSHI
; APPLICANT: ENDO, MASAHICO
; APPLICANT: KATO, IKUNOSHIN
; TITLE OF INVENTION: FUNGAL ANTIGENS AND PROCESS FOR PRODUCING THE SAME
; FILE REFERENCE: 1422-372P
; CURRENT APPLICATION NUMBER: US/09/262,856A
; CURRENT FILING DATE: 1999-03-04
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 721
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-262-856A-8

Alignment Scores:
Pred. No.: 0.000283 Length: 721
Score: 71.00 Matches: 13
Percent Similarity: 93.75% Conservative: 2
Best Local Similarity: 81.25% Mismatches: 1
Query Match: 48.97% Indels: 0
DB: 4 Gaps: 0

US-09-987-190-2 (1-30) x US-09-262-856A-8 (1-721)

Qy 13 AlaThrGluProTyrIleSerGlyGlnIleAenGluIle**TyrThr 28
Db 1 GCCATGAACCTATCATCATCAGACGAATGAACGAATTCATCACTACT 48

RESULT 19
5240847-7
; Patent No. 5240847
; APPLICANT: HECKL, KONRAD; SPEVAK, WALTER; OSTERMANN, ELINBORG;
; ZOPHEL, ANDREAS; KRYSZEK, EDELTHAUD; MAURER-FOCY, INGRID;
; WITCHE-CASTRANO, MARIA J.; STRATOWA, CHRISTIAN; HAUPTMANN, RUDOLF
; TITLE OF INVENTION: HUMAN MANGANESE SUPEROXIDE DISMUTASE
; (HMN-SOD)
; NUMBER OF SEQUENCES: 34
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/167,261
; FILING DATE: 11-MAR-1988
; SEQ ID NO.7:
; LENGTH: 99

Alignment Scores:
```

```

5240847-7
Alignment Scores:
Pred. No.: 0.000119 Length: 99
Score: 67.00 Matches: 12
Percent Similarity: 76.00% Conservative: 7
Best Local Similarity: 48.00% Mismatches: 6
Query Match: 46.21% Indels: 0
DB: 6 Gaps: 0

US-09-987-190-2 (1-30) x 5240847-7 (1-99)

Qy 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20
Db 4 AAGCAGCTTTGCCAGACTTCATACGACTACGCTGCTAGAACACACATCATGCT 63
21 GlnIleAenGluIle 25
Db 64 CAAATCATGCAATT 78

RESULT 20
US-09-091-097-5
; Sequence 5, Application US/09091097
; Patent No. 6432407
; GENERAL INFORMATION:
; APPLICANT: TAKESAKO, KAZUTOH
; APPLICANT: OKADO, TAKASHI
; APPLICANT: YAGIHARA, TOMOKO
; APPLICANT: KURODA, MASANOBU
; APPLICANT: ONISHI, YOSHIMI
; APPLICANT: KATO, IKUNOSHIN
; APPLICANT: AKIYAMA, KAZUO
; APPLICANT: YASUEDA, HIROSHI
; APPLICANT: YAMAGUCHI, HIDEYO
; TITLE OF INVENTION: ANTIGENIC PROTEIN ORIGINATING IN
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH, LLP
; STREET: PO BOX 747
; CITY: FALLS CHURCH
; STATE: VA
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/091,097
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: WEINER, MARC S.
; REGISTRATION NUMBER: 32,181
; REFERENCE/DOCKET NUMBER: 1422-0346P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-205-8000
; TELEFAX: 703-205-8050
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 728 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..618
US-09-091-097-5

Alignment Scores:
```



```

Db      37  TTTGAATTGCCAGCATTCGGTATCGTATGATGGCGTTGACCGCAGATCGACAAAGAA  96
OY      22  IleaenGuIle***TyRThr 28
          |||  ::|||
Db      97  ACGATGAACATTCACCACACG  117

RESULT 24
US-08-445-909A-16
; Sequence 16, Application US/08445909A
; Patent No. 5772996
; GENERAL INFORMATION:
; APPLICANT: ATKINSON, Antony
; TITLE OF INVENTION: Pharmaceutical Compositions
; TITLE OF INVENTION: Containing Superoxide Dismutase from Bacillus
; TITLE OF INVENTION: Stearothermophilus and Bacillus Caldotenax
; NUMBER OF SEQUENCES: 29
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/445,909A
; FILING DATE: 22-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/978,697
; FILING DATE: 02-FEB-1993
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1294 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 387..1001
; US-08-445-909A-16

Alignment Scores:
Pred. No.: 1,42 Length: 1294
Score: 53.00 Matches: 10
Percent Similarity: 55.56% Conservative: 5
Best Local Similarity: 37.04% Mismatches: 12
Query Match: 36.55% Indels: 0
DB: 1 Gaps: 0

99-987-190-2 (1-30) x US-08-445-909A-16 (1-1294)
OY      2  TyrSerLeuProGluIleuAspTyrgluPheSerAlaThrGluProTyrlIeSerGlyIn 21
          ::||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      393 TTTGAATTGCCAGCATTCGGTATCGTATGATGGCTTGACCGCAGATCGACAAAGAA  452
OY      22  IleaenGuIle***TyRThr 28
          |||  ::|||
Db      453 ACGATGAACATTCACCACACG  473

RESULT 25
US-08-894-772-4
; Sequence 4, Application US/08894772
; Patent No. 5861280
; GENERAL INFORMATION:
; APPLICANT: Lehmbeck, Jan
; TITLE OF INVENTION: Host Cell Expressing Reduced Levels
; TITLE OF INVENTION: Of A Metalloprotease And Methods Using The Host Cell In Protea
; TITLE OF INVENTION: Production
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 5861280 No. 5861280disk of No. 5861280th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York

```

```

STATE: New York
COUNTRY: U.S.A.
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/894,772
FILING DATE: 27-AUG-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Gregg, Valeta A.
REGISTRATION NUMBER: 35,127
REFERENCE/DOCKET NUMBER: 4300.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-867-0298
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 747 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-894-772-4

Alignment Scores:
Pred. No.: 2.04 Length: 747
Score: 50.50 Matches: 12
Percent Similarity: 53.33% Conservative: 4
Best Local Similarity: 40.00% Mismatches: 9
Query Match: 34.83% Indels: 5
DB: 2 Gaps: 2

US-09-987-190-2 (1-30) x US-08-894-772-4 (1-747)
Qy 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluPro-----TyrIle 18
Db 202 AAATTCATCATACCC-----TATCGCTTCACGCCACACACCCCGAGCTCTATATT 252
Qy 19 SerGlyGlnIleAsnGluIle**TyrThr 28
Db 253 GATGCGCTATCATCTACGCTTTCTTCACT 282

RESULT 26
US-09-207-844-4
Sequence 4, Application US/09207844
Patent No. 5968774
GENERAL INFORMATION:
APPLICANT: Lehbeck, Jan
TITLE OF INVENTION: Host Cell Expressing Reduced Levels
TITLE OF INVENTION: Of A Metalloprotease And Methods Using The Host Cell In Protein
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSER: No. 5968774o No. 5968774disk of No. 5968774th America, Inc.
STREET: 405 Lexington Avenue, Suite 6400
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/207,844

```

```

; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/894,772
; FILING DATE: 27-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Gregg, Valecia A.
; REGISTRATION NUMBER: 35,127
; REFERENCE/DOCKET NUMBER: 4300.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-867-0288
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 747 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; 9-207-844-4

Alignment Scores:
Pred. No.: 2 04 Length: 747
Score: 50.50 Matches: 12
Percent Similarity: 53.33% Conservative: 4
Best Local Similarity: 40.00% Mismatches: 9
Query Match: 34.83% Indels: 5
DB: 2 Gaps: 2

US-09-987-190-2 (1-30) x US-09-207-844-4 (1-747)
Qy 1 TySerleupProgluLeuApTyrgluPheserAlathrgluPro-----TyrIle 18
Db 202 AAATTCGAATACCC-----TATCGTTCAACGCCACACCCCGAGTCCATATT 252
Qy 19 SerGlyInIleangluile***TyrThr 28
Db 253 GATGCGTCTACTACTGACGCTTTCTACACT 282

RESULT 27
US-09-605-785-703
; Sequence 703, Application US/09605785
; Patent No. 6321716
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Scolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasar A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C16
; CURRENT APPLICATION NUMBER: US/09/605,785
; FILING DATE: 2000-06-27
; NUMBER OF SEQ ID NOS: 835
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 703
; LENGTH: 2904
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-605-785-703
```

```

Alignment Scores:
Pred. No.: 14.9 Length: 2904
Score: 50.00 Matches: 9
Percent Similarity: 65.00% Conservative: 4
Best Local Similarity: 45.00% Mismatches: 7
Query Match: 34.48% Indels: 0
DB: 4 Gaps: 0

US-09-987-190-2 (1-30) x US-09-605-785-703 (1-2904)
Qy 2 TySerleupProgluLeuApTyrgluPheserAlathrgluProTyrIleSerGlyIn 21
Db 40 TACCTCGCTGCTGCAATGACTGGGACACCAAGTCCCTGGCCCTTACTGGGACACCCAG 99

RESULT 28
US-09-020-956-110
; Sequence 110, Application US/09020956
; Patent No. 6261562
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS FO
; NUMBER OF SEQUENCES: 178
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/020,956
; FILING DATE: 09-FEB-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.427C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 110:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3410 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
US-09-020-956-110

Alignment Scores:
Pred. No.: 18.3 Length: 3410
Score: 50.00 Matches: 9
Percent Similarity: 65.00% Conservative: 4
Best Local Similarity: 45.00% Mismatches: 7
Query Match: 34.48% Indels: 0
DB: 4 Gaps: 0

US-09-987-190-2 (1-30) x US-09-020-956-110 (1-3410)
Qy 2 TySerleupProgluLeuApTyrgluPheserAlathrgluProTyrIleSerGlyIn 21
Db 809 TACCTCGCTGCTGCAATGACTGGGACACCAAGTCCCTGGCCCTTACTGGGACACCCAG 868

RESULT 29
```

```
US-09-030-607-110
; Sequence 110, Application US/09030607
; Patent No. 6262245
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS FO
; NUMBER OF SEQUENCES: 224
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/030.607
; FILING DATE: 25-FEB-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 110:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3410 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; US-09-030-607-110

Alignment Scores:
Pred. No.: 18.3 Length: 3410
Score: 50.00 Matches: 9
Percent Similarity: 65.00% Conservative: 4
Best Local Similarity: 45.00% Mismatches: 7
Query Match: 34.48% Indels: 0
Gaps: 0

9-987-190-2 (1-30) x US-09-030-607-110 (1-3410)

QY 2 TySerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGlyGln 21
Db 809 TACCTCCGCTGCGCATGACATGGAGACACCACTGCGCCCTACTGCGGACCCAG 868

RESULT 30
US-09-605-785-110
; Sequence 110, Application US/09605785
; Patent No. 6321716
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stoik, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
```

```
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C16
; CURRENT APPLICATION NUMBER: US/09/605.785
; CURRENT FILING DATE: 2000-06-27
; NUMBER OF SEQ ID NOS: 835
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 110
; LENGTH: 3410
; TYPE: DNA
; ORGANISM: Homo sapien
; US-09-605-785-110

Alignment Scores:
Pred. No.: 18.3 Length: 3410
Score: 50.00 Matches: 9
Percent Similarity: 65.00% Conservative: 4
Best Local Similarity: 45.00% Mismatches: 7
Query Match: 34.48% Indels: 0
Gaps: 0

US-09-987-190-2 (1-30) x US-09-605-785-110 (1-3410)

QY 2 TySerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGlyGln 21
Db 809 TACCTCCGCTGCGCATGACATGGAGACACCACTGCGCCCTACTGCGGACCCAG 868

RESULT 31
US-09-439-313-110
; Sequence 110, Application US/09439313
; Patent No. 6329505
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang Yugu
; APPLICANT: Reed, Steven G.
; APPLICANT: Kalos, Michael
; APPLICANT: Fanger, Gary
; APPLICANT: Retter, Mark
; APPLICANT: Solik, John
; APPLICANT: Day, Craig
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.427C9
; CURRENT APPLICATION NUMBER: US/09/439.313
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 575
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 110
; LENGTH: 3410
; TYPE: DNA
; ORGANISM: Homo sapien
; US-09-439-313-110

Alignment Scores:
Pred. No.: 18.3 Length: 3410
Score: 50.00 Matches: 9
Percent Similarity: 65.00% Conservative: 4
Best Local Similarity: 45.00% Mismatches: 7
Query Match: 34.48% Indels: 0
Gaps: 0

US-09-987-190-2 (1-30) x US-09-439-313-110 (1-3410)

QY 2 TySerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGlyGln 21
Db 809 TACCTCCGCTGCGCATGACATGGAGACACCACTGCGCCCTACTGCGGACCCAG 868
```



Db 809 TACCTCCTGCTGCATTGACTGGGACACGTCCTGCGCCCTTACTGCGACCCAG 868

RESULT 32

US-09-352-616A-110

Sequence 110, Application US/09352616A

Patent No. 6395278

GENERAL INFORMATION:

APPLICANT: Dillon, Davin C.

APPLICANT: Harlocker, Susan Louise

APPLICANT: Jiang, Yugu

APPLICANT: Xu, Jiangchun

APPLICANT: Mitcham, Jennifer Lynn

TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS

TITLE OF INVENTION: OF PROSTATE CANCER AND METHODS FOR THEIR USE

FILE REFERENCE: 210121.427C8

CURRENT APPLICATION NUMBER: US/09/352,616A

CURRENT FILING DATE: 1999-07-13

NUMBER OF SEQ ID NOS: 472

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 110

LENGTH: 3410

TYPE: DNA

ORGANISM: Homo sapien

US-09-352-616A-110

Alignment Scores:

Pred. No.: 18.3 Length: 3410

Score: 50.00 Matches: 9

Percent Similarity: 65.00% Conservative: 4

Best Local Similarity: 45.00% Mismatches: 7

Query Match: 34.48% Indels: 0

Gaps: 0

DB:

US-09-987-190-2 (1-30) x US-09-352-616A-110 (1-3410)

Qy 2 TySerleupProgluLeuAspTyrGluPheserAlaThrGluProTyrIleSerGlyGln 21

Db 809 TACCTCCTGCTGCATTGACTGGGACACGTCCTGCGCCCTTACTGCGACCCAG 868

RESULT 33

US-09-602-877A-100

Sequence 100, Application US/09602877A

Patent No. 6432707

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.

APPLICANT: Xu, Jiangchun

APPLICANT: Dillon, Davin C.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER

FILE REFERENCE: 210121.446C5

CURRENT APPLICATION NUMBER: US/09/602,877A

CURRENT FILING DATE: 2000-06-22

NUMBER OF SEQ ID NOS: 107

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 100

LENGTH: 3410

TYPE: DNA

ORGANISM: Homo sapien

US-09-602-877A-100

Alignment Scores:

Pred. No.: 18.3 Length: 3410

Score: 50.00 Matches: 9

Percent Similarity: 65.00% Conservative: 4

Best Local Similarity: 45.00% Mismatches: 7

Query Match: 34.48% Indels: 0

Gaps: 0

DB:

US-09-987-190-2 (1-30) x US-09-602-877A-100 (1-3410)

Qy 2 TySerleupProgluLeuAspTyrGluPheserAlaThrGluProTyrIleSerGlyGln 21

Db 809 TACCTCCTGCTGCATTGACTGGGACACGTCCTGCGCCCTTACTGCGACCCAG 868

RESULT 34

US-09-232-149A-110

Sequence 110, Application US/09232149A

Patent No. 6465611

GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun

APPLICANT: Dillon, Davin C.

APPLICANT: Mitcham, Jennifer Lynn

TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE

TITLE OF INVENTION: CANCER AND METHODS FOR THEIR USE

FILE REFERENCE: 210121.427C6

CURRENT APPLICATION NUMBER: US/09/232,149A

CURRENT FILING DATE: 1999-01-15

NUMBER OF SEQ ID NOS: 338

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 110

LENGTH: 3410

TYPE: DNA

ORGANISM: Homo sapien

US-09-232-149A-110

Alignment Scores:

Pred. No.: 18.3 Length: 3410

Score: 50.00 Matches: 9

Percent Similarity: 65.00% Conservative: 4

Best Local Similarity: 45.00% Mismatches: 7

Query Match: 34.48% Indels: 0

Gaps: 0

DB:

US-09-987-190-2 (1-30) x US-09-232-149A-110 (1-3410)

Qy 2 TySerleupProgluLeuAspTyrGluPheserAlaThrGluProTyrIleSerGlyGln 21

Db 809 TACCTCCTGCTGCATTGACTGGGACACGTCCTGCGCCCTTACTGCGACCCAG 868

RESULT 35

US-09-605-785-704

Sequence 704, Application US/09605785

Patent No. 6321716

GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun

APPLICANT: Dillon, Davin C.

APPLICANT: Mitcham, Jennifer L.

APPLICANT: Harlocker, Susan L.

APPLICANT: Jiang, Yugu

APPLICANT: Henderson, Robert A.

APPLICANT: Kalos, Michael D.

APPLICANT: Fanger, Gary R.

APPLICANT: Retter, Marc W.

APPLICANT: Stolk, John A.

APPLICANT: Day, Craig H.

APPLICANT: Vedvick, Thomas S.

APPLICANT: Carter, Darrick

APPLICANT: Li, Samuel

APPLICANT: Wang, Aijun

APPLICANT: Skeiky, Yasir A.W.

APPLICANT: Hepler, William

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER

FILE REFERENCE: 210121.427C16

CURRENT APPLICATION NUMBER: US/09/605,785

CURRENT FILING DATE: 2000-06-27

NUMBER OF SEQ ID NOS: 835

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 704

LENGTH: 4034

TYPE: DNA

ORGANISM: Homo sapiens

US-09-605-785-704

Alignment Scores:

Pred. No.: 22.8 Length: 4034

Score: 50.00 Matches: 9  
Percent Similarity: 65.00% Conservative: 4  
Best Local Similarity: 45.00% Mismatches: 7  
Query Match: 34.48% Indels: 0  
DB: 4 Gaps: 0

US-09-987-190-2 (1-30) x US-09-605-785-704 (1-4034)

Oy 2 TyrsleupProgluLeuAspTyrgluPheserAlaThrGluProTyrlleserGlyGln 21

Db 806 TACCTCCTGCTGCGATGACGACACAGTCCCTGCGCCCTACTGCGGACCCAG 865

## RESULT 36

US-09-221-017B-329/c  
Sequence 329, Application US/09221017B  
Patent No. 6444799

GENERAL INFORMATION:

APPLICANT: Ross, Bruce C.

TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF

NUMBER OF SEQUENCES: 1120

CORRESPONDENCE ADDRESS:

ADDRESSER: MORRISON & ROEBSTER

STREET: 755 PAGE MILL ROAD

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304-1018

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

OPERATING SYSTEM: Windows

SOFTWARE: FastSeq for Windows Version 2.0b

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/221.017B

FILING DATE: 23-DEC-1998

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: P1182

FILING DATE: 31-DEC-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: P11546

FILING DATE: 30-JAN-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: P2311

FILING DATE: 09-APR-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/AU98/01023

FILING DATE: 10-DEC-1998

ATTORNEY/AGENT INFORMATION:

NAME: Monroy, Gladys H

REGISTRATION NUMBER: 32,430

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-813-5600

TELEFAX: 650-494-0792

TELEX: 706141

INFORMATION FOR SEQ ID NO: 329:

SEQUENCE CHARACTERISTICS:

LENGTH: 4137 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: circular

MOLECULE TYPE: DNA (genomic)

HYPOTHEICAL: NO

ANTI-SENSE: UNKNOWN

ORIGINAL SOURCE:

ORGANISM: PORPHYROMONAS GINGIVALIS

FEATURE:

NAME/KEY: misc\_feature

LOCATION: 1...4137

US-09-221-017B-329

Alignment Scores:

Pred. No.: 23.6 Length: 4137  
Score: 50.00 Matches: 9  
Percent Similarity: 76.47% Conservative: 4  
Best Local Similarity: 52.94% Mismatches: 7  
Query Match: 34.48% Indels: 0  
DB: 4 Gaps: 0

US-09-987-190-2 (1-30) x US-09-221-017B-329 (1-4137)

Oy 5 ProgluLeuAspTyrgluPheserAlaThrGluProTyrlleserGlyGln 21

Db 3384 CCCGAATGAAATACAAAGCTGCTCAGATCCGTACATCTGACACCA 3334

## RESULT 37

US-09-605-785-702  
Sequence 702, Application US/09605785  
Patent No. 6321716

GENERAL INFORMATION:

APPLICANT: Xu, Jianshun

APPLICANT: Dillon, Davin C.

APPLICANT: Mitcham, Jennifer L.

APPLICANT: Harlocker, Susan L.

APPLICANT: Jiang, Yuqi

APPLICANT: Henderson, Robert A.

APPLICANT: Kalos, Michael D.

APPLICANT: Fanger, Gary R.

APPLICANT: Retter, Marc W.

APPLICANT: Stolk, John A.

APPLICANT: Day, Craig H.

APPLICANT: Vedvick, Thomas S.

APPLICANT: Carter, Darick

APPLICANT: Li, Samuel

APPLICANT: Wang, Aijun

APPLICANT: Hepler, William

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

FILE REFERENCE: 210121.427C16

CURRENT APPLICATION NUMBER: US/09/605.785

NUMBER OF SEQ ID NOS: 835

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 702

LENGTH: 4894

TYPE: DNA

ORGANISM: Homo sapiens

US-09-605-785-702

Alignment Scores:

Pred. No.: 29.3 Length: 4894

Score: 50.00 Matches: 9

Percent Similarity: 65.00% Conservative: 4

Best Local Similarity: 45.00% Mismatches: 7

Query Match: 34.48% Indels: 0

DB: 4 Gaps: 0

US-09-987-190-2 (1-30) x US-09-605-785-702 (1-4894)

Oy 2 TyrsleupProgluLeuAspTyrgluPheserAlaThrGluProTyrlleserGlyGln 21

Db 1665 TACCTCCTGCTGCGATGACGACACAGTCCCTGCGCCCTACTGCGGACCCAG 1724

## RESULT 38

US-09-605-785-705  
Sequence 705, Application US/09605785  
Patent No. 6321716

GENERAL INFORMATION:

APPLICANT: Xu, Jianshun

APPLICANT: Dillon, Davin C.

APPLICANT: Mitcham, Jennifer L.

APPLICANT: Harlocker, Susan L.

APPLICANT: Jiang, Yuqi

APPLICANT: Henderson, Robert A.

APPLICANT: Kalos, Michael D.  
APPLICANT: Fanger, Gary R.  
APPLICANT: Retter, Marc W.  
APPLICANT: Stolk, John A.  
APPLICANT: Day, Craig H.  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Carter, Derrick  
APPLICANT: Li, Samuel  
APPLICANT: Wang, Aijun  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Hepler, William  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
FILE REFERENCE: 210121.427C16  
CURRENT APPLICATION NUMBER: US/09/605,785  
CURRENT FILING DATE: 2000-06-27  
NUMBER OF SEQ ID NOS: 835  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 705  
LENGTH: 6976  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-605-785-705

Alignment Scores:  
Pred. No.: 46.6 Length: 6976  
Score: 50.00 Matches: 9  
Percent Similarity: 65.00% Conservative: 4  
Best Local Similarity: 45.00% Mismatches: 7  
Query Match: 34.48% Indels: 0  
DB: 4 Gaps: 0

US-09-987-190-2 (1-30) x US-09-605-785-705 (1-6976)

QY 2 TyrsenleuprogulneupgyluPheserlaThrguProTyrlleserGlyGln 21  
DB 1556 TACCTCCCTGCTGACCTGACTGAGACACCAAGTCCCTGAGCCCTTACTGAGCACCAG 1615

RESULT 39  
US-09-798-096-10/c  
Sequence 10, Application US/09798096  
Patent No. 6399378  
GENERAL INFORMATION:  
APPLICANT: Donna T. Ward  
APPLICANT: Andrew T. Watt  
TITLE OF INVENTION: ANTISENSE MODULATION OF RECOL2 EXPRESSION  
FILE REFERENCE: RUS-0207  
CURRENT APPLICATION NUMBER: US/09/798,096  
CURRENT FILING DATE: 2001-03-01  
NUMBER OF SEQ ID NOS: 89  
SEQ ID NO 10  
LENGTH: 99500  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
US-09-798-096-10

Alignment Scores:  
Pred. No.: 2.29e+03 Length: 99500  
Score: 49.00 Matches: 8  
Percent Similarity: 63.64% Conservative: 6  
Best Local Similarity: 36.36% Mismatches: 8  
Query Match: 33.79% Indels: 0  
DB: 4 Gaps: 0

US-09-987-190-2 (1-30) x US-09-798-096-10 (1-99500)

QY 4 leuprogulneupgyluPheserlaThrguProTyrlleserGlyGlnleasn 23  
DB 84101 CTGCTGACCTGACCTATGAGTTTCTTCACTGACCAACGATGGGTTCACCACTTAT 84042  
QY 24 Gluile 25  
:::||||

DB 84041 CAAATC 84036

RESULT 40  
US-09-735-935-3  
Sequence 3, Application US/09735935  
Patent No. 6420150  
GENERAL INFORMATION:  
APPLICANT: GUEGLER, Karl et al  
TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING  
TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN  
TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS,  
TITLE OF INVENTION: AND USES THEREOF  
FILE REFERENCE: CL000970  
CURRENT APPLICATION NUMBER: US/09/735,935  
CURRENT FILING DATE: 2000-12-14  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 5044  
TYPE: DNA  
ORGANISM: Human  
US-09-735-935-3

Alignment Scores:  
Pred. No.: 58.2 Length: 5044  
Score: 48.50 Matches: 10  
Percent Similarity: 66.67% Conservative: 6  
Best Local Similarity: 41.67% Mismatches: 7  
Query Match: 33.45% Indels: 1  
DB: 4 Gaps: 1

US-09-987-190-2 (1-30) x US-09-735-935-3 (1-5044)

QY 3 SerleuprogulneupgyluPheserlaThrguProTyrlleserGlyGln 21  
DB 275 TCCCTCCAAATCCPAACCTGATGATTTTCAGTCTGGACCATCTCAGAAATCCATG 334

QY 22 Ileasngluile 25  
DB 335 CTTTCAGAAATG 346

Search completed: April 8, 2003, 03:51:30  
Job time : 22.3578 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 9, 2003, 14:17:20 ; Search time 6.05505 Seconds  
(without alignments)  
302.901 Million cell updates/sec

Title: US-09-987-190-2  
Perfect score: 30  
Sequence: 1 KYSLPELDYEFSAPEYISQINEIXYTX 30

Scoring table:  
Gapop 60.0 , Gapext 60.0

Searched: 248812 seqs, 61136040 residues  
Word size: 0

Total number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database:

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubppaa/US08\_NEW\_PUB pep:\*
- 2: /cgn2\_6/ptodata/2/pubppaa/PCR\_NEW\_PUB pep:\*
- 3: /cgn2\_6/ptodata/2/pubppaa/US06\_NEW\_PUB pep:\*
- 4: /cgn2\_6/ptodata/2/pubppaa/US06\_PUBCOMB pep:\*
- 5: /cgn2\_6/ptodata/2/pubppaa/US07\_NEW\_PUB pep:\*
- 6: /cgn2\_6/ptodata/2/pubppaa/US07\_PUBCOMB pep:\*
- 7: /cgn2\_6/ptodata/2/pubppaa/PCUTUS\_PUBCOMB pep:\*
- 8: /cgn2\_6/ptodata/2/pubppaa/US08\_PUBCOMB pep:\*
- 9: /cgn2\_6/ptodata/2/pubppaa/US09\_NEW\_PUB pep:\*
- 10: /cgn2\_6/ptodata/2/pubppaa/US09\_PUBCOMB pep:\*
- 11: /cgn2\_6/ptodata/2/pubppaa/US10\_NEW\_PUB pep:\*
- 12: /cgn2\_6/ptodata/2/pubppaa/US10\_PUBCOMB pep:\*
- 13: /cgn2\_6/ptodata/2/pubppaa/US60\_NEW\_PUB pep:\*
- 14: /cgn2\_6/ptodata/2/pubppaa/US60\_PUBCOMB pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	27	90.0	30	10	US-09-987-190-2	Sequence 2, Appl
2	6	20.0	50	10	US-09-864-761-43011	Sequence 4011, Appl
3	6	20.0	65	12	US-10-078-929-24	Sequence 24, Appl
4	6	20.0	81	9	US-09-981-876-166	Sequence 166, Appl
5	6	20.0	81	9	US-09-148-545-166	Sequence 166, Appl
6	20.0	98	12	US-10-078-929-26	Sequence 26, Appl	
7	6	20.0	137	10	US-09-840-787-49	Sequence 49, Appl
8	6	20.0	146	10	US-09-925-300-975	Sequence 975, Appl
9	6	20.0	188	10	US-09-987-190-6	Sequence 6, Appl
10	6	20.0	200	9	US-09-738-626-62727	Sequence 62727, Appl
11	6	20.0	200	10	US-09-818-564-2	Sequence 2, Appl
12	6	20.0	1536	9	US-09-909-567B-47	Sequence 47, Appl
13	5	16.7	26	9	US-09-813-153-211	Sequence 211, Appl
14	5	16.7	33	10	US-09-864-761-48159	Sequence 48159, Appl
15	5	16.7	39	10	US-09-810-310-2	Sequence 2, Appl
16	5	16.7	39	10	US-09-810-310-9	Sequence 2, Appl
17	5	16.7	39	10	US-09-864-761-34370	Sequence 34370, Appl
18	5	16.7	52	9	US-09-764-904-63	Sequence 63, Appl
19	5	16.7	52	9	US-10-091-548-63	Sequence 63, Appl

20	5	16.7	52	10	US-09-764-860-581	Sequence 581, App
21	5	16.7	65	10	US-09-864-761-34543	Sequence 34543, A
22	5	16.7	76	10	US-09-867-550-424	Sequence 424, App
23	5	16.7	77	10	US-09-864-761-47808	Sequence 47808, A
24	5	16.7	88	10	US-09-764-887-257	Sequence 257, App
25	5	16.7	91	10	US-09-864-761-40079	Sequence 40079, A
26	5	16.7	97	9	US-10-101-464A-651	Sequence 651, App
27	5	16.7	105	10	US-09-764-877-1863	Sequence 1863, App
28	5	16.7	111	12	US-10-052-817-6	Sequence 6, App1
29	5	16.7	127	10	US-09-846-590B-2	Sequence 2, App1
30	5	16.7	132	10	US-09-164-615-13	Sequence 13, App1
31	5	16.7	133	9	US-10-101-464A-485	Sequence 485, App
32	5	16.7	136	9	US-10-147-454-1	Sequence 1, App1
33	5	16.7	137	10	US-10-147-454-3	Sequence 3, App1
34	5	16.7	137	10	US-09-291-332-1	Sequence 1, App1
35	5	16.7	139	9	US-10-147-454-2	Sequence 2, App1
36	5	16.7	140	9	US-09-738-626-3519	Sequence 3519, App
37	5	16.7	152	12	US-10-094-214-4	Sequence 4, App1
38	5	16.7	157	9	US-09-791-932-73	Sequence 73, App1
39	5	16.7	159	9	US-10-112-793-16	Sequence 16, App1
40	5	16.7	159	10	US-10-101-464A-650	Sequence 650, App
41	5	16.7	159	10	US-09-884-909-6	Sequence 6, App1
42	5	16.7	159	10	US-09-884-909-6	Sequence 6, App1
43	5	16.7	159	10	US-09-800-908-15	Sequence 15, App1
44	5	16.7	162	10	US-09-925-297-616	Sequence 616, App1
45	5	16.7	169	10	US-09-775-925-29	Sequence 29, App1
46	5	16.7	169	10	US-09-847-519A-14	Sequence 14, App1
47	5	16.7	170	12	US-10-109-885-3	Sequence 3, App1
48	5	16.7	171	9	US-09-764-868-1095	Sequence 1095, App
49	5	16.7	191	10	US-09-878-454A-2	Sequence 2, App1
50	5	16.7	191	12	US-10-109-885-2	Sequence 2, App1
51	5	16.7	196	9	US-09-738-626-5550	Sequence 5550, App
52	5	16.7	198	9	US-09-727-855B-7	Sequence 7, App1
53	5	16.7	211	9	US-10-270-877-46	Sequence 46, App1
54	5	16.7	211	9	US-10-270-877-46	Sequence 46, App1
55	5	16.7	213	10	US-09-815-242-13061	Sequence 13061, App
56	5	16.7	225	10	US-09-905-810-1	Sequence 1, App1
57	5	16.7	239	10	US-09-741-669-438	Sequence 438, App
58	5	16.7	246	9	US-10-037-677-2	Sequence 2, App1
59	5	16.7	249	9	US-09-880-748-1160	Sequence 1160, App
60	5	16.7	249	10	US-09-905-810-2	Sequence 2, App1
61	5	16.7	249	10	US-09-782-980-44	Sequence 44, App1
62	5	16.7	250	12	US-10-052-817-4	Sequence 4, App1
63	5	16.7	251	9	US-09-880-748-1688	Sequence 1688, App
64	5	16.7	251	10	US-09-821-831-4	Sequence 4, App1
65	5	16.7	252	9	US-09-738-626-6778	Sequence 6778, App
66	5	16.7	262	9	US-10-101-464A-649	Sequence 649, App
67	5	16.7	273	10	US-09-742-454A-2	Sequence 2, App1
68	5	16.7	273	10	US-09-883-777-2	Sequence 2, App1
69	5	16.7	275	12	US-09-846-808-8	Sequence 8, App1
70	5	16.7	275	12	US-10-059-964-28	Sequence 28, App1
71	5	16.7	279	12	US-10-059-964-60	Sequence 60, App1
72	5	16.7	281	12	US-10-059-964-58	Sequence 58, App1
73	5	16.7	282	9	US-10-051-643-134	Sequence 134, App
74	5	16.7	282	9	US-09-880-505-134	Sequence 134, App
75	5	16.7	284	10	US-09-894-998-41	Sequence 41, App1
76	5	16.7	285	9	US-10-101-464A-692	Sequence 692, App
77	5	16.7	285	10	US-09-846-808-9	Sequence 9, App1
78	5	16.7	285	10	US-10-059-964-30	Sequence 30, App1
79	5	16.7	289	10	US-09-886-868-24	Sequence 24, App1
80	5	16.7	291	10	US-09-846-808-10	Sequence 10, App1
81	5	16.7	291	12	US-10-059-964-18	Sequence 18, App1
82	5	16.7	293	10	US-09-846-808-13	Sequence 13, App1
83	5	16.7	293	12	US-10-062-624-40	Sequence 40, App1
84	5	16.7	293	12	US-10-059-964-44	Sequence 44, App1
85	5	16.7	293	12	US-10-059-964-54	Sequence 54, App1
86	5	16.7	295	10	US-09-846-808-5	Sequence 5, App1
87	5	16.7	295	12	US-10-059-964-22	Sequence 22, App1
88	5	16.7	296	12	US-10-059-964-56	Sequence 56, App1
89	5	16.7	297	9	US-10-051-643-177	Sequence 177, App
90	5	16.7	297	9	US-09-880-505-177	Sequence 177, App
91	5	16.7	297	12	US-10-059-964-14	Sequence 14, App1
92	5	16.7	298	9	US-09-813-153-212	Sequence 212, App

93	5	16.7	298	10	US-09-846-808-11	Sequence 11, Appl	166	5	16.7	471	9	US-09-880-505-172	Sequence 172, App
94	5	16.7	298	12	US-10-059-964-52	Sequence 52, Appl	167	5	16.7	479	9	US-10-108-605-57	Sequence 57, Appl
95	5	16.7	301	9	US-10-028-072-166	Sequence 166, App	168	5	16.7	484	10	US-09-846-5908-4	Sequence 4, Appl1
96	5	16.7	301	9	US-10-123-049-166	Sequence 166, App	169	5	16.7	491	9	US-10-237-684-34	Sequence 34, Appl
97	5	16.7	301	9	US-10-123-904-166	Sequence 166, App	170	5	16.7	491	9	US-10-230-163-34	Sequence 34, Appl
98	5	16.7	301	9	US-10-140-470-166	Sequence 166, App	171	5	16.7	491	9	US-10-218-631-34	Sequence 34, Appl
99	5	16.7	301	9	US-10-175-746-166	Sequence 166, App	172	5	16.7	491	9	US-10-230-338-34	Sequence 34, Appl
100	5	16.7	301	9	US-10-176-918-166	Sequence 166, App	173	5	16.7	491	9	US-10-230-414-34	Sequence 34, Appl
101	5	16.7	301	9	US-10-176-921-166	Sequence 166, App	174	5	16.7	493	10	US-09-742-684-12	Sequence 12, Appl
102	5	16.7	301	9	US-10-137-865-166	Sequence 166, App	175	5	16.7	500	10	US-09-925-300-1362	Sequence 1362, App
103	5	16.7	301	9	US-10-140-474-166	Sequence 166, App	176	5	16.7	504	9	US-09-905-370-1	Sequence 1, Appl1
104	5	16.7	301	9	US-10-142-431-166	Sequence 166, App	177	5	16.7	510	9	US-10-029-180-94	Sequence 94, Appl
105	5	16.7	301	9	US-10-143-114-166	Sequence 166, App	178	5	16.7	511	10	US-09-796-202-17	Sequence 17, Appl
106	5	16.7	301	9	US-10-140-002-166	Sequence 166, App	179	5	16.7	513	12	US-10-143-002-2	Sequence 2, Appl1
107	5	16.7	301	9	US-10-142-419-166	Sequence 166, App	180	5	16.7	516	9	US-10-026-741-48	Sequence 48, Appl
108	5	16.7	301	9	US-10-123-262-166	Sequence 166, App	181	5	16.7	519	10	US-09-756-551A-8	Sequence 8, Appl1
109	5	16.7	301	9	US-10-142-423-166	Sequence 166, App	182	5	16.7	521	10	US-09-815-242-11855	Sequence 11855, A
110	5	16.7	301	9	US-10-121-050-166	Sequence 166, App	183	5	16.7	530	9	US-10-028-072-130	Sequence 130, App
111	5	16.7	301	9	US-10-141-755-166	Sequence 166, App	184	5	16.7	530	9	US-10-121-049-130	Sequence 130, App
112	5	16.7	301	9	US-10-143-033-166	Sequence 166, App	185	5	16.7	530	9	US-10-123-904-130	Sequence 130, App
113	5	16.7	302	9	US-09-736-457-806	Sequence 806, App	186	5	16.7	530	9	US-10-140-470-130	Sequence 130, App
114	5	16.7	302	9	US-09-902-941-806	Sequence 806, App	187	5	16.7	530	9	US-10-175-746-130	Sequence 130, App
115	5	16.7	302	9	US-09-849-626-806	Sequence 806, App	188	5	16.7	530	9	US-10-176-918-130	Sequence 130, App
116	5	16.7	302	9	US-10-017-754-806	Sequence 806, App	189	5	16.7	530	9	US-10-176-921-130	Sequence 130, App
117	5	16.7	316	9	US-10-101-464A-941	Sequence 941, App	190	5	16.7	530	9	US-10-137-665-130	Sequence 130, App
118	5	16.7	321	9	US-10-012-507-1	Sequence 1, Appl1	191	5	16.7	530	9	US-10-140-474-130	Sequence 130, App
119	5	16.7	331	9	US-09-738-626-4960	Sequence 4960, App	192	5	16.7	530	9	US-10-142-431-130	Sequence 130, App
120	5	16.7	335	10	US-09-738-626-4330	Sequence 4330, App	193	5	16.7	530	9	US-10-143-114-130	Sequence 130, App
121	5	16.7	338	10	US-09-815-242-13457	Sequence 13457, A	194	5	16.7	530	9	US-10-140-002-130	Sequence 130, App
122	5	16.7	338	10	US-09-815-242-13645	Sequence 13645, A	195	5	16.7	530	9	US-10-142-419-130	Sequence 130, App
123	5	16.7	339	10	US-09-802-371-2	Sequence 2, Appl1	196	5	16.7	530	9	US-10-123-662-130	Sequence 130, App
124	5	16.7	341	10	US-09-916-790-18	Sequence 18, Appl	197	5	16.7	530	9	US-10-142-423-130	Sequence 130, App
125	5	16.7	369	9	US-09-738-626-5658	Sequence 5658, App	198	5	16.7	530	9	US-10-121-050-130	Sequence 130, App
126	5	16.7	371	9	US-09-870-759-107	Sequence 107, App	199	5	16.7	530	9	US-10-141-755-130	Sequence 130, App
127	5	16.7	386	9	US-10-101-464A-802	Sequence 902, App	200	5	16.7	530	9	US-10-143-032-130	Sequence 130, App
128	5	16.7	394	9	US-09-736-457-805	Sequence 805, App	201	5	16.7	563	10	US-09-789-561-83	Sequence 83, Appl
129	5	16.7	394	9	US-09-736-457-827	Sequence 827, App	202	5	16.7	566	10	US-09-829-482-5	Sequence 5, Appl1
130	5	16.7	394	9	US-09-902-941-805	Sequence 805, App	203	5	16.7	598	10	US-09-875-811-12	Sequence 12, Appl
131	5	16.7	394	9	US-09-902-941-827	Sequence 827, App	204	5	16.7	605	10	US-09-875-811-8	Sequence 8, Appl1
132	5	16.7	394	9	US-09-849-626-805	Sequence 805, App	205	5	16.7	615	10	US-09-771-838A-2	Sequence 2, Appl1
133	5	16.7	394	9	US-09-849-626-827	Sequence 827, App	206	5	16.7	615	10	US-09-815-242-5211	Sequence 5211, App
134	5	16.7	394	9	US-10-017-754-805	Sequence 805, App	207	5	16.7	621	10	US-09-875-811-4	Sequence 4, Appl1
135	5	16.7	394	9	US-10-017-754-827	Sequence 827, App	208	5	16.7	645	9	US-10-108-605-285	Sequence 285, App
136	5	16.7	396	9	US-09-987-107-33	Sequence 33, Appl	209	5	16.7	649	10	US-09-815-242-11306	Sequence 1306, A
137	5	16.7	396	9	US-09-802-640-16	Sequence 16, Appl	210	5	16.7	650	10	US-09-815-242-13667	Sequence 13667, A
138	5	16.7	396	10	US-09-800-729-207	Sequence 207, App	211	5	16.7	656	10	US-09-875-811-10	Sequence 10, Appl
139	5	16.7	401	9	US-09-902-941-1817	Sequence 1917, App	212	5	16.7	666	10	US-09-842-446-2	Sequence 2, Appl1
140	5	16.7	401	9	US-09-849-626-1917	Sequence 1917, App	213	5	16.7	663	10	US-09-875-811-6	Sequence 6, Appl1
141	5	16.7	401	9	US-10-017-754-1917	Sequence 1917, App	214	5	16.7	665	9	US-09-533-029-76	Sequence 76, Appl
142	5	16.7	406	10	US-09-924-256A-25	Sequence 25, Appl	215	5	16.7	679	10	US-09-875-811-2	Sequence 2, Appl1
143	5	16.7	419	10	US-09-741-669-415	Sequence 415, Appl	216	5	16.7	682	10	US-09-771-838A-3	Sequence 3, Appl1
144	5	16.7	421	10	US-09-742-954-14	Sequence 14, Appl	217	5	16.7	682	10	US-09-079-892-1	Sequence 1, Appl1
145	5	16.7	425	9	US-09-813-398-32	Sequence 32, Appl	218	5	16.7	715	9	US-10-101-664A-924	Sequence 924, App
146	5	16.7	425	10	US-09-748-537-14	Sequence 14, Appl	219	5	16.7	742	9	US-10-108-605-203	Sequence 203, App
147	5	16.7	425	10	US-09-821-831-2	Sequence 26, Appl	220	5	16.7	774	9	US-10-162-706-5	Sequence 5, Appl1
148	5	16.7	427	8	US-08-681-219-26	Sequence 26, Appl	221	5	16.7	790	10	US-09-801-368-184	Sequence 184, App
149	5	16.7	427	10	US-09-826-212-5	Sequence 5, Appl1	222	5	16.7	797	9	US-10-066-443-2	Sequence 2, Appl1
150	5	16.7	427	10	US-09-826-212-5	Sequence 13, Appl	223	5	16.7	797	10	US-09-066-236-2	Sequence 2, Appl1
151	5	16.7	427	10	US-09-748-537-13	Sequence 7, Appl1	224	5	16.7	798	9	US-10-066-443-5	Sequence 5, Appl1
152	5	16.7	432	9	US-09-738-626-4678	Sequence 4678, App	225	5	16.7	798	10	US-09-900-236-5	Sequence 5, Appl1
153	5	16.7	433	10	US-09-801-368-226	Sequence 226, App	226	5	16.7	807	10	US-09-847-046-2	Sequence 2, Appl1
154	5	16.7	445	9	US-09-738-626-6951	Sequence 6951, App	227	5	16.7	838	10	US-09-894-998-45	Sequence 45, Appl
155	5	16.7	445	9	US-10-067-974-12	Sequence 12, Appl	228	5	16.7	853	9	US-10-003-035-33	Sequence 33, Appl
156	5	16.7	445	9	US-10-226-136-20	Sequence 20, Appl	229	5	16.7	856	10	US-09-476-242-1	Sequence 1, Appl1
157	5	16.7	449	9	US-09-738-626-6150	Sequence 6150, App	230	5	16.7	861	10	US-10-026-741-103	Sequence 103, App
158	5	16.7	450	10	US-09-896-720-2	Sequence 2, Appl1	231	5	16.7	861	10	US-09-815-242-11152	Sequence 11152, A
159	5	16.7	455	9	US-09-756-854-4	Sequence 4, Appl1	232	5	16.7	869	10	US-09-815-242-110623	Sequence 21, Appl
160	5	16.7	455	9	US-10-041-574-4	Sequence 121, App	233	5	16.7	871	10	US-09-886-468-21	Sequence 21, Appl
161	5	16.7	455	9	US-10-051-643-121	Sequence 121, App	234	5	16.7	893	9	US-10-014-436-4	Sequence 4, Appl1
162	5	16.7	455	9	US-09-880-505-121	Sequence 121, App	235	5	16.7	908	10	US-09-815-242-11847	Sequence 11847, A
163	5	16.7	457	9	US-10-124-601-2	Sequence 2, Appl1	236	5	16.7	911	10	US-09-828-423-4	Sequence 4, Appl1
164	5	16.7	459	9	US-09-746-660A-44	Sequence 44, Appl	237	5	16.7	950	10	US-09-815-242-11271	Sequence 11271, A
165	5	16.7	471	9	US-10-051-643-172	Sequence 172, App	238	5	16.7	951	10	US-09-815-242-10465	Sequence 10465, A

239	5	16.7	951	10	US-09-815-242-13853	Sequence 13853, A	312	4	13.3	11	9	US-09-861-661-28	Sequence 28, Appl
240	5	16.7	960	8	US-08-910-866A-18	Sequence 18, Appl	313	4	13.3	11	10	US-09-758-128-16	Sequence 16, Appl
241	5	16.7	969	10	US-09-118-276-2	Sequence 2, Appl	314	4	13.3	11	10	US-09-758-128-19	Sequence 19, Appl
242	5	16.7	980	10	US-09-118-276-11	Sequence 11, Appl	315	4	13.3	11	10	US-09-758-128-21	Sequence 22, Appl
243	5	16.7	1050	10	US-09-866-562-38	Sequence 38, Appl	316	4	13.3	11	10	US-09-758-128-25	Sequence 25, Appl
244	5	16.7	1101	9	US-10-003-035-53	Sequence 53, Appl	317	4	13.3	11	10	US-09-758-128-28	Sequence 28, Appl
245	5	16.7	1186	9	US-10-003-035-55	Sequence 55, Appl	318	4	13.3	12	10	US-09-982-172-201	Sequence 201, App
246	5	16.7	1198	9	US-09-975-119-405	Sequence 405, Appl	319	4	13.3	12	10	US-09-982-172-249	Sequence 249, App
247	5	16.7	1211	10	US-09-801-574-60	Sequence 60, Appl	320	4	13.3	13	9	US-09-764-868-1241	Sequence 1241, Ap
248	5	16.7	1285	10	US-09-925-301-1394	Sequence 1394, Ap	321	4	13.3	13	9	US-09-955-999-121	Sequence 121, App
249	5	16.7	1422	8	US-08-424-5508-86	Sequence 86, Appl	322	4	13.3	13	10	US-09-950-313-41	Sequence 41, Appl
250	5	16.7	1474	10	US-09-873-403-5	Sequence 5, Appl	323	4	13.3	14	9	US-09-229-551A-16	Sequence 16, Appl
251	5	16.7	1474	12	US-10-052-817-2	Sequence 2, Appl	324	4	13.3	14	10	US-09-879-936-7	Sequence 7, Appl
252	5	16.7	1531	10	US-09-876-889-347	Sequence 347, App	325	4	13.3	15	1	US-08-464-363-13	Sequence 19, Appl
253	5	16.7	1531	10	US-09-998-598-2593	Sequence 2593, Ap	326	4	13.3	16	10	US-09-770-940-14	Sequence 14, Appl
254	5	16.7	1713	10	US-09-737-149-27	Sequence 27, Appl	327	4	13.3	16	10	US-09-150-947B-7	Sequence 7, Appl
255	5	16.7	1938	9	US-10-014-436-2	Sequence 2, Appl	328	4	13.3	17	9	US-09-986-480-407	Sequence 407, App
256	5	16.7	2025	10	US-09-815-242-5703	Sequence 5703, Ap	329	4	13.3	18	10	US-09-864-761-47951	Sequence 47951, A
257	5	16.7	2045	9	US-09-736-968A-109	Sequence 109, App	330	4	13.3	18	10	US-09-789-561-161	Sequence 161, App
258	5	16.7	2045	10	US-09-736-969A-95	Sequence 95, Appl	331	4	13.3	19	9	US-09-963-339-10	Sequence 10, Appl
259	5	16.7	2045	9	US-09-736-960-92	Sequence 92, Appl	332	4	13.3	19	9	US-10-062-710-113	Sequence 113, App
260	5	16.7	2047	9	US-09-736-968A-2	Sequence 2, Appl	333	4	13.3	19	10	US-09-799-576A-13	Sequence 13, Appl
261	5	16.7	2227	9	US-10-104-966-12	Sequence 12, Appl	334	4	13.3	19	10	US-09-799-540-13	Sequence 13, Appl
262	5	16.7	2227	9	US-10-135-988-2	Sequence 2, Appl	335	4	13.3	19	10	US-09-799-608-13	Sequence 13, Appl
263	5	16.7	2227	9	US-10-135-988-4	Sequence 4, Appl	336	4	13.3	19	12	US-10-059-964-63	Sequence 63, Appl
264	5	16.7	2227	9	US-10-135-988-6	Sequence 6, Appl	337	4	13.3	20	9	US-10-038-612-47	Sequence 47, Appl
265	5	16.7	2227	10	US-09-929-955-12	Sequence 12, Appl	338	4	13.3	20	9	US-09-592-982-1	Sequence 1, Appl
266	5	16.7	2383	10	US-09-912-020-302	Sequence 302, App	339	4	13.3	20	10	US-09-815-626-10	Sequence 10, Appl
267	5	16.7	2437	10	US-09-815-242-583A	Sequence 583A, Ap	340	4	13.3	21	9	US-10-038-612-117	Sequence 117, Appl
268	5	16.7	2448	10	US-09-815-242-11870	Sequence 11870, A	341	4	13.3	22	9	US-09-865-889-117	Sequence 117, App
269	5	16.7	2923	10	US-09-788-711A-4	Sequence 4, Appl	342	4	13.3	22	9	US-10-097-065-577	Sequence 577, App
270	5	16.7	2956	10	US-09-788-711A-2	Sequence 2, Appl	343	4	13.3	22	9	US-10-099-574A-117	Sequence 117, App
271	5	16.7	3014	10	US-09-737-149-25	Sequence 25, Appl	344	4	13.3	22	10	US-09-953-510-23	Sequence 23, Appl
272	5	16.7	3034	10	US-09-737-149-25	Sequence 25, Appl	345	4	13.3	22	10	US-09-867-852-155	Sequence 155, App
273	5	16.7	3034	10	US-09-737-149-30	Sequence 30, Appl	346	4	13.3	23	9	US-10-037-296-6	Sequence 6, Appl
274	5	16.7	3158	10	US-09-815-242-12611	Sequence 12611, A	347	4	13.3	23	9	US-09-782-816A-1	Sequence 1, Appl
275	5	16.7	3354	9	US-10-160-758-11	Sequence 11, Appl	348	4	13.3	23	10	US-09-827-949-30	Sequence 30, Appl
276	5	16.7	3354	9	US-10-160-758-12	Sequence 12, Appl	349	4	13.3	23	10	US-09-864-761-37623	Sequence 37623, A
277	5	16.7	4349	9	US-10-160-758-15	Sequence 15, Appl	350	4	13.3	23	10	US-09-864-761-46097	Sequence 46097, A
278	5	16.7	5701	10	US-09-864-761-37319	Sequence 37319, A	351	4	13.3	23	12	US-10-071-751-4	Sequence 4, Appl
279	5	16.7	6281	10	US-09-815-242-12396	Sequence 12396, A	352	4	13.3	24	9	US-09-864-480-224	Sequence 224, Appl
280	5	16.7	26926	9	US-09-759-5088-2	Sequence 2, Appl	353	4	13.3	24	10	US-09-864-761-34770	Sequence 34770, A
281	4	13.3	4	9	US-09-264-516A-9	Sequence 9, Appl	354	4	13.3	25	9	US-09-945-182-6	Sequence 6, Appl
282	4	13.3	4	10	US-09-834-765-737	Sequence 737, App	355	4	13.3	25	9	US-09-945-182-8	Sequence 8, Appl
283	4	13.3	5	10	US-09-799-118-8	Sequence 8, Appl	356	4	13.3	25	9	US-09-774-639-311	Sequence 311, App
284	4	13.3	6	9	US-09-809-391-622	Sequence 622, App	357	4	13.3	25	10	US-09-969-730-238	Sequence 238, App
285	4	13.3	6	12	US-10-156-820-67	Sequence 67, App	358	4	13.3	25	10	US-09-864-761-46556	Sequence 46556, A
286	4	13.3	7	9	US-09-782-397-20	Sequence 20, Appl	359	4	13.3	25	10	US-09-815-248-32	Sequence 22, Appl
287	4	13.3	7	9	US-09-782-397-22	Sequence 22, Appl	360	4	13.3	25	10	US-09-815-248-36	Sequence 36, Appl
288	4	13.3	7	9	US-09-782-397-25	Sequence 25, Appl	361	4	13.3	26	1	US-08-464-363-9	Sequence 9, Appl
289	4	13.3	7	9	US-09-782-397-27	Sequence 27, Appl	362	4	13.3	26	10	US-09-864-761-34796	Sequence 34796, A
290	4	13.3	9	9	US-10-062-710-133	Sequence 133, App	363	4	13.3	26	10	US-09-864-761-38523	Sequence 38523, A
291	4	13.3	9	10	US-09-834-765-156	Sequence 156, App	364	4	13.3	26	10	US-09-734-417-7	Sequence 7, Appl
292	4	13.3	9	10	US-09-753-831-7	Sequence 7, Appl	365	4	13.3	27	9	US-10-091-504-747	Sequence 747, App
293	4	13.3	10	9	US-10-025-222A-31	Sequence 31, Appl	366	4	13.3	27	10	US-09-864-761-46828	Sequence 46828, A
294	4	13.3	10	10	US-09-056-160B-91	Sequence 91, Appl	367	4	13.3	27	10	US-09-864-761-47610	Sequence 47610, A
295	4	13.3	10	10	US-09-834-765-83	Sequence 83, Appl	368	4	13.3	27	10	US-09-764-869-747	Sequence 747, App
296	4	13.3	10	10	US-09-834-765-96	Sequence 96, Appl	369	4	13.3	28	10	US-09-770-940-15	Sequence 15, Appl
297	4	13.3	11	9	US-09-758-426-16	Sequence 16, Appl	370	4	13.3	29	10	US-09-864-761-48447	Sequence 48447, A
298	4	13.3	11	9	US-09-758-426-19	Sequence 19, Appl	371	4	13.3	30	9	US-10-187-088-3	Sequence 3, Appl
299	4	13.3	11	9	US-09-758-426-22	Sequence 22, Appl	372	4	13.3	30	10	US-09-864-761-44274	Sequence 44274, A
300	4	13.3	11	9	US-09-758-426-25	Sequence 25, Appl	373	4	13.3	30	10	US-09-822-295-16	Sequence 16, Appl
301	4	13.3	11	9	US-09-758-426-28	Sequence 28, Appl	374	4	13.3	30	10	US-09-925-300-1135	Sequence 1135, Ap
302	4	13.3	11	9	US-09-758-198-16	Sequence 16, Appl	375	4	13.3	31	9	US-09-764-868-1258	Sequence 1258, Ap
303	4	13.3	11	9	US-09-758-198-19	Sequence 19, Appl	376	4	13.3	31	9	US-09-955-899-138	Sequence 138, App
304	4	13.3	11	9	US-09-758-198-22	Sequence 22, Appl	377	4	13.3	31	10	US-09-730-772-19	Sequence 19, Appl
305	4	13.3	11	9	US-09-758-198-25	Sequence 25, Appl	378	4	13.3	31	10	US-09-730-772-23	Sequence 23, Appl
306	4	13.3	11	9	US-09-758-198-28	Sequence 28, Appl	379	4	13.3	31	10	US-09-735-849-19	Sequence 19, Appl
307	4	13.3	11	9	US-10-166-158-1	Sequence 1, Appl	380	4	13.3	31	10	US-09-735-849-23	Sequence 23, Appl
308	4	13.3	11	9	US-09-861-661-16	Sequence 16, Appl	381	4	13.3	31	10	US-09-732-091-9	Sequence 9, Appl
309	4	13.3	11	9	US-09-861-661-19	Sequence 19, Appl	382	4	13.3	32	9	US-09-964-130-65	Sequence 65, Appl
310	4	13.3	11	9	US-09-861-661-22	Sequence 22, Appl	383	4	13.3	32	10	US-09-864-761-43735	Sequence 43735, A
311	4	13.3	11	9	US-09-861-661-25	Sequence 25, Appl	384	4	13.3	32	10	US-09-864-761-44091	Sequence 44091, A

385	4	13.3	32	10	US-09-815-248-34	Sequence 34, Appl	458	4	13.3	50	9	US-10-043-452-15	Sequence 15, Appl
386	4	13.3	33	10	US-09-864-761-34780	Sequence 34780, A	459	4	13.3	50	9	US-09-796-692-147	Sequence 747, App
387	4	13.3	33	10	US-09-864-761-42715	Sequence 42715, A	460	4	13.3	50	10	US-09-864-761-38709	Sequence 38709, A
388	4	13.3	34	9	US-09-776-724A-278	Sequence 278, App	461	4	13.3	50	10	US-09-864-761-40402	Sequence 40402, A
389	4	13.3	34	10	US-09-776-724A-278	Sequence 278, App	462	4	13.3	50	10	US-09-864-761-41965	Sequence 41965, A
390	4	13.3	35	9	US-10-223-047-24	Sequence 24, Appl	463	4	13.3	50	10	US-09-864-761-42100	Sequence 42100, A
391	4	13.3	35	9	US-09-855-698-8	Sequence 8, Appl1	464	4	13.3	50	10	US-09-864-761-41845	Sequence 41845, A
392	4	13.3	35	10	US-09-776-490-8	Sequence 8, Appl1	465	4	13.3	50	12	US-10-062-254-1172	Sequence 172, App
393	4	13.3	35	10	US-09-776-491-8	Sequence 8, Appl1	466	4	13.3	51	9	US-09-974-879-224	Sequence 224, App
394	4	13.3	35	10	US-09-150-947B-8	Sequence 8, Appl1	467	4	13.3	51	9	US-10-091-504-1204	Sequence 1204, App
395	4	13.3	35	10	US-09-864-761-41265	Sequence 41265, A	468	4	13.3	51	10	US-09-764-869-1204	Sequence 1204, App
396	4	13.3	35	10	US-09-864-761-4100	Sequence 4100, A	469	4	13.3	51	10	US-09-764-877-1474	Sequence 1474, App
397	4	13.3	36	9	US-09-983-802-182	Sequence 182, App	470	4	13.3	51	10	US-09-764-877-1935	Sequence 1935, App
398	4	13.3	36	9	US-10-050-704-293	Sequence 293, App	471	4	13.3	52	9	US-09-782-816A-51	Sequence 51, Appl
399	4	13.3	36	9	US-09-864-761-45243	Sequence 45243, A	472	4	13.3	52	9	US-10-083-357-1128	Sequence 1128, App
400	4	13.3	37	9	US-09-738-626-6768	Sequence 6768, App	473	4	13.3	52	9	US-10-091-504-739	Sequence 739, App
401	4	13.3	37	9	US-10-091-504-665	Sequence 665, App	474	4	13.3	52	10	US-09-864-761-35037	Sequence 35037, A
402	4	13.3	37	10	US-09-864-761-37614	Sequence 37614, A	475	4	13.3	52	10	US-09-864-761-37309	Sequence 37309, A
403	4	13.3	37	10	US-09-764-869-665	Sequence 665, App	476	4	13.3	52	10	US-09-864-761-44482	Sequence 44482, A
404	4	13.3	38	9	US-09-776-724A-249	Sequence 249, App	477	4	13.3	52	10	US-09-864-761-46407	Sequence 46407, A
405	4	13.3	38	9	US-10-196-107A-65	Sequence 65, Appl	478	4	13.3	53	9	US-09-764-869-739	Sequence 739, App
406	4	13.3	38	10	US-09-374-671-65	Sequence 65, Appl	479	4	13.3	53	9	US-09-738-626-6792	Sequence 6792, App
407	4	13.3	38	10	US-09-864-761-43271	Sequence 43271, A	480	4	13.3	53	9	US-10-007-521-48	Sequence 48, Appl
408	4	13.3	38	10	US-09-764-898-171	Sequence 171, App	481	4	13.3	53	9	US-10-092-154-751	Sequence 751, App
409	4	13.3	38	10	US-09-764-877-2002	Sequence 2002, App	482	4	13.3	53	10	US-09-864-761-43425	Sequence 43425, A
410	4	13.3	39	9	US-09-983-802-572	Sequence 572, App	483	4	13.3	53	10	US-09-864-761-43322	Sequence 43622, A
411	4	13.3	39	12	US-10-042-417-22	Sequence 22, Appl	484	4	13.3	53	10	US-09-764-847-751	Sequence 751, App
412	4	13.3	40	9	US-09-994-595-119	Sequence 119, App	485	4	13.3	53	10	US-09-764-877-1983	Sequence 1983, App
413	4	13.3	40	9	US-10-000-256A-179	Sequence 179, App	486	4	13.3	53	10	US-09-925-300-1523	Sequence 1523, App
414	4	13.3	40	10	US-09-864-761-41532	Sequence 41532, A	487	4	13.3	54	9	US-10-016-634A-130	Sequence 130, App
415	4	13.3	41	9	US-09-966-422B-16	Sequence 16, Appl	488	4	13.3	54	10	US-09-864-761-33909	Sequence 33909, A
416	4	13.3	41	10	US-09-864-761-40304	Sequence 40304, A	489	4	13.3	54	10	US-09-864-761-39351	Sequence 39351, A
417	4	13.3	41	10	US-09-864-761-42107	Sequence 42107, A	490	4	13.3	54	10	US-09-764-898-273	Sequence 273, App
418	4	13.3	41	10	US-09-815-248-35	Sequence 35, Appl	491	4	13.3	55	9	US-09-796-692-1378	Sequence 1378, App
419	4	13.3	42	10	US-09-903-452-12	Sequence 12, Appl	492	4	13.3	55	9	US-09-796-692-2059	Sequence 2059, App
420	4	13.3	42	10	US-09-864-761-43393	Sequence 43393, A	493	4	13.3	55	9	US-09-796-692-2350	Sequence 2350, App
421	4	13.3	42	10	US-09-864-761-48554	Sequence 48554, A	494	4	13.3	55	9	US-10-091-572-293	Sequence 293, App
422	4	13.3	42	10	US-09-764-877-1278	Sequence 1278, App	495	4	13.3	55	10	US-09-864-761-42808	Sequence 42908, A
423	4	13.3	43	10	US-09-205-658-195	Sequence 195, App	496	4	13.3	55	10	US-09-864-761-44051	Sequence 44051, A
424	4	13.3	43	10	US-09-864-761-39366	Sequence 39366, A	497	4	13.3	55	10	US-09-864-761-48670	Sequence 48670, A
425	4	13.3	43	10	US-09-864-761-39856	Sequence 39856, A	498	4	13.3	55	10	US-09-815-242-23	Sequence 13031, A
426	4	13.3	43	10	US-09-864-761-11903	Sequence 41903, A	499	4	13.3	55	10	US-09-815-248-22	Sequence 23, Appl
427	4	13.3	43	10	US-09-864-761-42201	Sequence 42201, A	500	4	13.3	56	10	US-09-867-550-1844	Sequence 1844, App
428	4	13.3	43	10	US-09-864-761-42573	Sequence 42573, A	501	4	13.3	57	1	US-08-722-570-4	Sequence 4, Appl1
429	4	13.3	43	10	US-09-864-761-42822	Sequence 42822, A	502	4	13.3	57	10	US-09-864-761-38394	Sequence 38394, A
430	4	13.3	43	10	US-09-864-761-43463	Sequence 43463, A	503	4	13.3	57	10	US-09-864-761-38928	Sequence 38928, A
431	4	13.3	43	10	US-09-864-761-45936	Sequence 45936, A	504	4	13.3	57	10	US-09-925-297-891	Sequence 891, App
432	4	13.3	43	10	US-09-839-884-7	Sequence 7, Appl1	505	4	13.3	57	10	US-09-925-297-891	Sequence 3, Appl1
433	4	13.3	44	9	US-09-989-919-94	Sequence 94, Appl1	506	4	13.3	57	10	US-09-764-877-1396	Sequence 1956, App
434	4	13.3	44	9	US-09-983-802-359	Sequence 359, App	507	4	13.3	58	10	US-09-925-299-1300	Sequence 1300, App
435	4	13.3	44	9	US-09-969-730-163	Sequence 163, App	508	4	13.3	58	10	US-09-864-761-36106	Sequence 36106, A
436	4	13.3	44	10	US-09-864-761-35723	Sequence 35723, A	509	4	13.3	58	10	US-09-864-761-36816	Sequence 36816, A
437	4	13.3	44	10	US-09-864-761-38584	Sequence 38584, A	510	4	13.3	58	10	US-09-925-299-11300	Sequence 1300, App
438	4	13.3	44	10	US-09-864-761-11074	Sequence 41074, A	511	4	13.3	59	9	US-10-081-572-385	Sequence 355, App
439	4	13.3	44	10	US-09-864-761-45946	Sequence 45946, A	512	4	13.3	59	10	US-09-827-761-431	Sequence 41, Appl
440	4	13.3	44	10	US-09-815-242-12843	Sequence 12843, A	513	4	13.3	59	10	US-09-864-761-36759	Sequence 36759, A
441	4	13.3	45	9	US-09-774-639-184	Sequence 184, App	514	4	13.3	59	10	US-09-864-761-47674	Sequence 47674, A
442	4	13.3	45	9	US-10-092-154-865	Sequence 865, App	515	4	13.3	59	10	US-09-815-242-5539	Sequence 5539, App
443	4	13.3	45	10	US-09-864-761-35002	Sequence 35002, A	516	4	13.3	59	10	US-09-815-242-12823	Sequence 12823, A
444	4	13.3	45	10	US-09-764-847-865	Sequence 865, App	517	4	13.3	59	10	US-10-046-561-17	Sequence 12827, A
445	4	13.3	46	9	US-10-002-344A-268	Sequence 268, App	518	4	13.3	60	9	US-10-007-521-70	Sequence 17, Appl
446	4	13.3	46	10	US-09-864-761-44167	Sequence 44167, A	519	4	13.3	60	9	US-09-864-761-47405	Sequence 47405, A
447	4	13.3	46	10	US-09-864-761-48378	Sequence 48378, A	520	4	13.3	60	10	US-09-864-761-4705	Sequence 47405, A
448	4	13.3	46	10	US-09-815-248-32	Sequence 32, Appl	521	4	13.3	61	9	US-09-981-876-138	Sequence 138, App
449	4	13.3	46	10	US-09-948-080-8	Sequence 8, Appl1	522	4	13.3	61	9	US-10-001-857-138	Sequence 138, App
450	4	13.3	47	10	US-09-864-761-37953	Sequence 37953, A	523	4	13.3	61	9	US-09-148-945-138	Sequence 138, App
451	4	13.3	47	10	US-09-864-761-46971	Sequence 46971, A	524	4	13.3	61	10	US-09-864-761-42382	Sequence 42382, A
452	4	13.3	48	10	US-09-864-761-48902	Sequence 48902, A	525	4	13.3	62	10	US-09-925-302-837	Sequence 837, App
453	4	13.3	48	10	US-09-764-877-1403	Sequence 1403, App	526	4	13.3	62	10	US-09-864-761-38482	Sequence 38482, A
454	4	13.3	49	10	US-09-864-761-35680	Sequence 35680, A	527	4	13.3	62	10	US-09-864-761-30034	Sequence 30034, A
455	4	13.3	49	10	US-09-864-761-47455	Sequence 47455, A	528	4	13.3	63	9	US-10-083-357-1168	Sequence 1168, App
456	4	13.3	49	10	US-09-815-248-28	Sequence 28, Appl	529	4	13.3	63	10	US-09-954-737-10	Sequence 10, Appl
457	4	13.3	49	10	US-09-948-080-1	Sequence 1, Appl1	530	4	13.3	64	9	US-10-000-256A-199	Sequence 199, App



531	4	13.3	64	10	US-09-864-761-38549	Sequence 38549, A	604	4	13.3	80	10	US-09-789-561-112	Sequence 112, App
532	4	13.3	64	10	US-09-864-761-43264	Sequence 43264, A	605	4	13.3	80	10	US-09-746-491-44	Sequence 44, Appl
533	4	13.3	64	10	US-09-864-761-46904	Sequence 46904, A	606	4	13.3	80	10	US-09-764-875-1226	Sequence 1226, Ap
534	4	13.3	64	10	US-09-734-417-15	Sequence 15, Appl	607	4	13.3	81	9	US-09-796-692-2291	Sequence 2291, Ap
535	4	13.3	65	9	US-09-738-626-4258	Sequence 4258, Ap	608	4	13.3	81	10	US-09-864-761-34288	Sequence 34288, A
536	4	13.3	65	9	US-09-983-802-538	Sequence 538, App	609	4	13.3	82	10	US-09-729-674-2275	Sequence 275, App
537	4	13.3	65	10	US-09-864-761-38878	Sequence 38878, A	610	4	13.3	82	10	US-09-864-761-36650	Sequence 36650, A
538	4	13.3	65	10	US-09-911-8264-20	Sequence 20, Appl	611	4	13.3	82	10	US-09-864-761-43086	Sequence 43086, A
539	4	13.3	66	9	US-10-092-154-502	Sequence 502, App	612	4	13.3	83	9	US-10-091-504-1007	Sequence 1007, Ap
540	4	13.3	66	10	US-09-939-980-485	Sequence 485, App	613	4	13.3	83	10	US-09-764-869-11007	Sequence 11007, Ap
541	4	13.3	67	10	US-09-764-847-502	Sequence 502, App	614	4	13.3	83	10	US-09-754-9974-116	Sequence 116, Appl
542	4	13.3	67	9	US-09-738-626-6646	Sequence 6646, App	615	4	13.3	84	10	US-09-864-761-37887	Sequence 37887, A
543	4	13.3	68	9	US-09-764-872-284	Sequence 284, App	616	4	13.3	84	10	US-09-864-761-44669	Sequence 44669, A
544	4	13.3	69	9	US-09-843-676-17	Sequence 17, Appl	617	4	13.3	84	10	US-09-867-550-596	Sequence 596, App
545	4	13.3	69	9	US-09-766-253-17	Sequence 17, Appl	618	4	13.3	84	10	US-09-764-877-11603	Sequence 1603, Ap
546	4	13.3	69	9	US-09-438-486-17	Sequence 17, Appl	619	4	13.3	86	10	US-09-764-883-438	Sequence 438, App
547	4	13.3	69	9	US-10-053-758-17	Sequence 17, Appl	620	4	13.3	87	10	US-09-867-550-648	Sequence 648, App
548	4	13.3	69	9	US-10-012-542-198	Sequence 198, App	621	4	13.3	87	10	US-09-815-248-4	Sequence 4, Appl1
549	4	13.3	69	9	US-10-034-295-17	Sequence 17, Appl	622	4	13.3	87	10	US-09-815-248-13	Sequence 13, Appl
550	4	13.3	69	9	US-09-809-391-606	Sequence 606, App	623	4	13.3	88	9	US-10-091-483-170	Sequence 170, App
551	4	13.3	69	9	US-10-092-154-689	Sequence 689, App	624	4	13.3	88	10	US-09-905-243-64	Sequence 64, Appl
552	4	13.3	69	9	US-10-054-611-17	Sequence 17, Appl	625	4	13.3	88	10	US-09-764-846-1170	Sequence 170, App
553	4	13.3	69	10	US-09-764-847-689	Sequence 689, App	626	4	13.3	89	9	US-09-738-626-4439	Sequence 4439, Ap
554	4	13.3	70	8	US-08-902-572-20	Sequence 20, Appl	627	4	13.3	89	9	US-09-764-872-459	Sequence 459, App
555	4	13.3	70	9	US-10-011-445-73	Sequence 73, Appl	628	4	13.3	89	9	US-10-079-854-135	Sequence 135, App
556	4	13.3	70	9	US-10-091-504-654	Sequence 654, App	629	4	13.3	89	9	US-10-102-627-49	Sequence 49, Appl
557	4	13.3	70	10	US-09-864-761-33329	Sequence 33329, A	630	4	13.3	89	10	US-09-864-761-34477	Sequence 135, App
558	4	13.3	70	10	US-09-864-761-35986	Sequence 35986, A	631	4	13.3	89	10	US-09-764-878-1135	Sequence 135, App
559	4	13.3	70	10	US-09-864-761-45440	Sequence 45440, A	632	4	13.3	89	10	US-09-764-860-472	Sequence 472, App
560	4	13.3	70	10	US-09-764-869-654	Sequence 654, App	633	4	13.3	89	10	US-10-083-357-762	Sequence 762, App
561	4	13.3	70	10	US-09-920-552-95	Sequence 95, Appl	634	4	13.3	90	10	US-09-205-658-247	Sequence 247, App
562	4	13.3	71	10	US-09-864-761-35705	Sequence 35705, A	635	4	13.3	90	10	US-09-764-877-11488	Sequence 1488, Ap
563	4	13.3	71	10	US-09-864-761-36820	Sequence 36820, A	636	4	13.3	91	9	US-09-989-920-239	Sequence 239, App
564	4	13.3	71	10	US-09-864-761-37078	Sequence 37078, A	637	4	13.3	91	10	US-09-925-301-924	Sequence 924, App
565	4	13.3	71	10	US-09-864-761-45725	Sequence 45725, A	638	4	13.3	91	10	US-09-867-550-418	Sequence 418, App
566	4	13.3	71	10	US-09-864-761-46025	Sequence 46025, A	639	4	13.3	92	9	US-09-989-920-225	Sequence 225, App
567	4	13.3	71	10	US-09-815-242-4926	Sequence 4926, Ap	640	4	13.3	92	9	US-10-091-572-414	Sequence 414, App
568	4	13.3	71	10	US-09-867-550-1168	Sequence 1168, App	641	4	13.3	92	10	US-09-927-112-16	Sequence 16, Appl
569	4	13.3	71	10	US-09-764-877-1957	Sequence 1957, Ap	642	4	13.3	93	9	US-10-134-641-2	Sequence 2, Appl1
570	4	13.3	72	9	US-09-945-182-11	Sequence 11, Appl	643	4	13.3	93	9	US-10-083-357-782	Sequence 782, App
571	4	13.3	72	9	US-09-738-626-4708	Sequence 4708, Ap	644	4	13.3	94	9	US-10-091-572-371	Sequence 371, App
572	4	13.3	72	9	US-09-974-879-164	Sequence 164, App	645	4	13.3	94	10	US-09-764-853-716	Sequence 716, App
573	4	13.3	72	10	US-09-827-949-17	Sequence 17, Appl	646	4	13.3	95	10	US-10-076-785-44	Sequence 44, Appl
574	4	13.3	72	10	US-09-866-073-17	Sequence 17, Appl	647	4	13.3	95	10	US-09-822-635-8	Sequence 8, Appl1
575	4	13.3	73	9	US-10-007-2804-175	Sequence 175, App	648	4	13.3	95	10	US-09-864-761-49105	Sequence 49105, A
576	4	13.3	73	10	US-09-925-302-543	Sequence 543, App	649	4	13.3	95	10	US-09-800-971-10	Sequence 10, Appl
577	4	13.3	73	10	US-09-864-761-35171	Sequence 35171, A	650	4	13.3	95	10	US-09-927-112-9	Sequence 9, Appl1
578	4	13.3	73	10	US-09-864-761-48326	Sequence 48326, A	651	4	13.3	95	10	US-09-908-664-12	Sequence 12, Appl
579	4	13.3	73	10	US-09-867-550-946	Sequence 946, App	652	4	13.3	95	10	US-09-908-664-21	Sequence 21, Appl
580	4	13.3	73	10	US-09-529-063-46	Sequence 46, Appl	653	4	13.3	95	10	US-09-925-306-1335	Sequence 1335, Ap
581	4	13.3	73	10	US-09-948-080-17	Sequence 17, Appl	654	4	13.3	96	1	US-08-464-363-39	Sequence 39, Appl
582	4	13.3	74	9	US-09-738-626-6931	Sequence 6931, Ap	655	4	13.3	96	1	US-09-809-391-473	Sequence 473, App
583	4	13.3	74	9	US-10-214-188-10	Sequence 10, Appl	656	4	13.3	96	9	US-09-809-357-999	Sequence 999, App
584	4	13.3	74	10	US-09-864-761-36961	Sequence 36961, A	657	4	13.3	96	9	US-10-083-357-999	Sequence 999, App
585	4	13.3	75	10	US-09-864-761-48605	Sequence 48605, A	658	4	13.3	96	9	US-10-092-154-792	Sequence 792, App
586	4	13.3	75	10	US-09-833-067-9	Sequence 9, Appl1	659	4	13.3	96	9	US-10-091-504-1214	Sequence 1214, App
587	4	13.3	75	10	US-09-815-242-10918	Sequence 10918, A	660	4	13.3	96	10	US-09-864-761-47502	Sequence 47502, A
588	4	13.3	76	9	US-10-083-357-703	Sequence 703, App	661	4	13.3	96	10	US-09-833-067-11	Sequence 11, Appl
589	4	13.3	76	10	US-09-864-761-36343	Sequence 36343, A	662	4	13.3	96	10	US-09-764-865-1214	Sequence 1214, Ap
590	4	13.3	76	10	US-09-864-761-36363	Sequence 36363, A	663	4	13.3	97	9	US-09-764-884-792	Sequence 792, App
591	4	13.3	76	10	US-09-864-761-39363	Sequence 39363, A	664	4	13.3	97	9	US-09-776-724-241	Sequence 241, App
592	4	13.3	77	9	US-10-091-504-1073	Sequence 1073, Ap	665	4	13.3	97	10	US-09-864-761-34195	Sequence 34195, A
593	4	13.3	77	10	US-09-864-761-38096	Sequence 38096, Ap	666	4	13.3	98	10	US-09-925-300-1380	Sequence 1380, App
594	4	13.3	77	10	US-09-764-869-1073	Sequence 1073, Ap	667	4	13.3	98	10	US-09-789-561-1860	Sequence 1860, App
595	4	13.3	78	9	US-10-102-806-672	Sequence 672, App	668	4	13.3	98	10	US-09-867-550-442	Sequence 442, App
596	4	13.3	78	10	US-09-764-887-271	Sequence 271, App	669	4	13.3	98	10	US-09-910-151-11	Sequence 11, Appl
597	4	13.3	78	10	US-09-667-550-1780	Sequence 1780, Ap	670	4	13.3	99	9	US-09-738-626-4164	Sequence 4164, App
598	4	13.3	79	10	US-09-864-761-47317	Sequence 47317, A	671	4	13.3	99	9	US-09-738-626-4651	Sequence 4651, Ap
599	4	13.3	79	10	US-09-864-761-48849	Sequence 48849, A	672	4	13.3	99	10	US-10-083-357-827	Sequence 827, App
600	4	13.3	79	10	US-09-734-417-19	Sequence 19, Appl	673	4	13.3	99	10	US-09-941-999-12	Sequence 2, Appl1
601	4	13.3	80	9	US-10-001-873-31	Sequence 31, Appl	674	4	13.3	100	9	US-09-864-877-1608	Sequence 1608, Ap
602	4	13.3	80	9	US-10-091-572-378	Sequence 378, App	675	4	13.3	100	10	US-09-854-133-198	Sequence 198, App
603	4	13.3	80	10	US-09-864-761-45728	Sequence 45728, A	676	4	13.3	101	9	US-10-004-381-1	Sequence 1, Appl1

677	4	13.3	101	9	US-10-080-980-13	Sequence 13, Appl	750	4	13.3	113	9	US-09-899-044-84	Sequence 84, Appl
678	4	13.3	101	9	US-10-097-065-576	Sequence 576, Appl	751	4	13.3	113	9	US-09-994-595-157	Sequence 157, Appl
679	4	13.3	101	10	US-09-840-787-37	Sequence 37, Appl	752	4	13.3	113	9	US-10-101-464A-579	Sequence 579, Appl
680	4	13.3	101	10	US-09-841-132-516	Sequence 516, Appl	753	4	13.3	113	10	US-09-294-121A-84	Sequence 84, Appl
681	4	13.3	101	10	US-09-854-731-2	Sequence 2, Appl	754	4	13.3	113	10	US-09-815-248-2	Sequence 2, Appl
682	4	13.3	101	10	US-09-751-1008-89	Sequence 89, Appl	755	4	13.3	113	10	US-09-815-248-9	Sequence 9, Appl
683	4	13.3	102	9	US-10-001-857-189	Sequence 189, Appl	756	4	13.3	113	10	US-09-899-082A-84	Sequence 84, Appl
684	4	13.3	102	10	US-09-925-301-1512	Sequence 1512, Ap	757	4	13.3	113	10	US-09-764-877-1547	Sequence 1547, Ap
685	4	13.3	102	10	US-09-841-132-586	Sequence 586, Appl	758	4	13.3	113	10	US-09-263-959-318	Sequence 318, Appl
686	4	13.3	103	9	US-09-974-298-4	Sequence 4, Appl	759	4	13.3	114	9	US-10-170-102-12	Sequence 12, Appl
687	4	13.3	103	10	US-09-764-877-1964	Sequence 1964, Ap	760	4	13.3	114	9	US-09-759-1308-198	Sequence 198, Appl
688	4	13.3	104	9	US-09-898-659-39	Sequence 39, Appl	761	4	13.3	114	9	US-10-227-884-52	Sequence 52, Appl
689	4	13.3	104	9	US-09-746-783-206	Sequence 206, Appl	762	4	13.3	114	9	US-10-230-163-52	Sequence 52, Appl
690	4	13.3	104	10	US-09-893-737-64	Sequence 64, Appl	763	4	13.3	114	9	US-10-218-631-52	Sequence 52, Appl
691	4	13.3	104	10	US-09-751-1008-72	Sequence 72, Appl	764	4	13.3	114	9	US-10-230-338-52	Sequence 52, Appl
692	4	13.3	105	9	US-10-187-088-9	Sequence 9, Appl	765	4	13.3	114	9	US-10-230-414-52	Sequence 52, Appl
693	4	13.3	105	9	US-10-043-487-314	Sequence 314, App	766	4	13.3	114	10	US-09-860-351-5	Sequence 5, Appl
694	4	13.3	105	9	US-10-125-540-357	Sequence 357, App	767	4	13.3	114	10	US-09-764-864-903	Sequence 903, Appl
695	4	13.3	105	10	US-09-764-870-357	Sequence 357, App	768	4	13.3	114	10	US-09-925-300-1551	Sequence 1551, Ap
696	4	13.3	105	10	US-09-864-761-39486	Sequence 39486, A	769	4	13.3	115	10	US-09-864-761-37760	Sequence 37760, A
697	4	13.3	105	10	US-09-925-301-1635	Sequence 1635, Ap	770	4	13.3	116	9	US-10-101-464A-96	Sequence 96, Appl
698	4	13.3	105	10	US-09-867-550-1366	Sequence 1366, Ap	771	4	13.3	116	10	US-09-867-550-146	Sequence 146, Appl
699	4	13.3	105	10	US-09-229-200A-10	Sequence 10, Appl	772	4	13.3	117	10	US-09-829-482-26	Sequence 26, Appl
700	4	13.3	106	9	US-09-764-872-417	Sequence 417, App	773	4	13.3	118	9	US-09-925-299-1167	Sequence 1167, App
701	4	13.3	106	9	US-10-229-394-1	Sequence 1, Appl	774	4	13.3	118	10	US-09-925-299-1167	Sequence 1167, Ap
702	4	13.3	106	10	US-09-815-242-11082	Sequence 11082, A	775	4	13.3	118	10	US-09-972-086-10	Sequence 10, Appl
703	4	13.3	106	10	US-09-751-1008-94	Sequence 94, Appl	776	4	13.3	119	9	US-09-738-626-4000	Sequence 4000, Appl
704	4	13.3	106	10	US-09-925-300-1776	Sequence 1776, Ap	777	4	13.3	119	9	US-09-738-626-4291	Sequence 4291, Ap
705	4	13.3	107	9	US-10-164-359-5	Sequence 5, Appl	778	4	13.3	119	9	US-09-925-299-908	Sequence 908, Appl
706	4	13.3	107	9	US-09-867-550-1154	Sequence 1154, Ap	779	4	13.3	119	9	US-09-925-299-908	Sequence 908, Appl
707	4	13.3	107	10	US-09-925-300-1874	Sequence 1874, Ap	780	4	13.3	119	10	US-09-908-871-82	Sequence 82, Appl
708	4	13.3	108	9	US-09-769-145-1	Sequence 1, Appl	781	4	13.3	119	10	US-09-925-299-908	Sequence 908, Appl
709	4	13.3	108	9	US-09-769-145-2	Sequence 2, Appl	782	4	13.3	119	10	US-09-925-299-908	Sequence 908, Appl
710	4	13.3	108	9	US-09-769-145-3	Sequence 3, Appl	783	4	13.3	120	9	US-09-925-299-908	Sequence 13, Appl
711	4	13.3	108	9	US-09-774-639-190	Sequence 190, App	784	4	13.3	120	9	US-09-989-293A-199	Sequence 199, App
712	4	13.3	108	9	US-09-778-026-3	Sequence 3, Appl	785	4	13.3	120	9	US-09-989-735-199	Sequence 199, App
713	4	13.3	108	9	US-09-778-026-4	Sequence 4, Appl	786	4	13.3	120	9	US-09-989-735-199	Sequence 199, App
714	4	13.3	108	9	US-09-778-026-5	Sequence 5, Appl	787	4	13.3	120	9	US-09-989-730-199	Sequence 199, App
715	4	13.3	108	9	US-09-969-730-151	Sequence 151, App	788	4	13.3	120	9	US-09-991-181-199	Sequence 199, App
716	4	13.3	108	10	US-09-864-761-34967	Sequence 34967, A	789	4	13.3	120	9	US-09-991-181-199	Sequence 199, App
717	4	13.3	108	10	US-09-234-395-6	Sequence 6, Appl	790	4	13.3	120	9	US-09-989-734-199	Sequence 199, App
718	4	13.3	108	10	US-09-234-395-7	Sequence 7, Appl	791	4	13.3	120	9	US-09-989-734-199	Sequence 199, App
719	4	13.3	108	10	US-09-234-395-8	Sequence 8, Appl	792	4	13.3	120	9	US-10-028-072-434	Sequence 434, App
720	4	13.3	108	10	US-09-751-1008-65	Sequence 65, Appl	793	4	13.3	120	9	US-09-989-735-199	Sequence 199, App
721	4	13.3	108	10	US-09-305-928-6	Sequence 6, Appl	794	4	13.3	120	9	US-09-993-667-199	Sequence 199, App
722	4	13.3	108	10	US-09-305-928-7	Sequence 7, Appl	795	4	13.3	120	9	US-10-121-049-434	Sequence 434, App
723	4	13.3	108	10	US-09-305-928-8	Sequence 8, Appl	796	4	13.3	120	9	US-10-123-904-434	Sequence 434, App
724	4	13.3	108	12	US-10-006-982-1	Sequence 1, Appl	797	4	13.3	120	9	US-10-140-470-434	Sequence 434, App
725	4	13.3	108	12	US-10-006-982-2	Sequence 2, Appl	798	4	13.3	120	9	US-09-990-438-199	Sequence 199, App
726	4	13.3	108	12	US-10-006-982-3	Sequence 3, Appl	799	4	13.3	120	9	US-09-990-438-199	Sequence 199, App
727	4	13.3	109	1	US-08-464-363-6	Sequence 6, Appl	800	4	13.3	120	9	US-09-997-666-199	Sequence 199, App
728	4	13.3	109	9	US-10-141-908-8	Sequence 8, Appl	801	4	13.3	120	9	US-10-175-746-434	Sequence 434, App
729	4	13.3	109	10	US-09-864-761-40024	Sequence 40024, A	802	4	13.3	120	9	US-10-176-518-434	Sequence 434, App
730	4	13.3	110	10	US-09-864-761-44550	Sequence 44550, A	803	4	13.3	120	9	US-10-176-518-434	Sequence 434, App
731	4	13.3	110	10	US-09-734-569-98	Sequence 98, Appl	804	4	13.3	120	9	US-10-176-518-434	Sequence 434, App
732	4	13.3	110	10	US-09-764-860-529	Sequence 529, Appl	805	4	13.3	120	9	US-09-990-711-199	Sequence 199, App
733	4	13.3	111	8	US-08-424-5508-69	Sequence 69, Appl	806	4	13.3	120	9	US-10-137-665-434	Sequence 434, App
734	4	13.3	111	9	US-09-965-528-9	Sequence 9, Appl	807	4	13.3	120	9	US-10-140-474-434	Sequence 434, App
735	4	13.3	111	9	US-09-738-626-3654	Sequence 3654, Ap	808	4	13.3	120	9	US-10-142-431-434	Sequence 434, App
736	4	13.3	111	9	US-10-103-196-22	Sequence 22, Appl	809	4	13.3	120	9	US-10-143-114-434	Sequence 434, App
737	4	13.3	111	10	US-09-864-761-37734	Sequence 37734, A	810	4	13.3	120	9	US-09-989-726-199	Sequence 199, App
738	4	13.3	112	9	US-09-974-879-188	Sequence 188, Appl	811	4	13.3	120	9	US-10-140-002-434	Sequence 434, App
739	4	13.3	112	10	US-09-864-761-33779	Sequence 33779, A	812	4	13.3	120	9	US-09-990-437-199	Sequence 199, App
740	4	13.3	112	10	US-09-864-761-40044	Sequence 40044, A	813	4	13.3	120	9	US-09-998-156-199	Sequence 199, App
741	4	13.3	112	10	US-09-731-672-338	Sequence 338, App	814	4	13.3	120	9	US-10-142-419-434	Sequence 434, App
742	4	13.3	113	9	US-09-899-302-84	Sequence 84, Appl	815	4	13.3	120	9	US-09-991-157-199	Sequence 199, App
743	4	13.3	113	9	US-09-851-138-70	Sequence 70, Appl	816	4	13.3	120	9	US-09-991-172-199	Sequence 199, App
744	4	13.3	113	9	US-09-851-138-72	Sequence 72, Appl	817	4	13.3	120	9	US-09-997-514-199	Sequence 199, App
745	4	13.3	113	9	US-09-851-138-74	Sequence 74, Appl	818	4	13.3	120	9	US-09-997-573-199	Sequence 199, App
746	4	13.3	113	9	US-09-851-138-78	Sequence 78, Appl	819	4	13.3	120	9	US-10-123-662-434	Sequence 434, App
747	4	13.3	113	9	US-09-851-138-80	Sequence 80, Appl	820	4	13.3	120	9	US-10-142-423-434	Sequence 434, App
748	4	13.3	113	9	US-09-899-046-146	Sequence 146, App	821	4	13.3	120	9	US-09-990-443-199	Sequence 199, App
749	4	13.3	113	9	US-09-878-281-146	Sequence 146, App	822	4	13.3	120	9	US-09-990-726-199	Sequence 199, App

823	4	13.3	120	9	US-09-997-559-199	Sequence 199, App	836	4	13.3	120	9	US-10-060-845-8	Sequence 8, Appli
824	4	13.3	120	9	US-09-997-601-199	Sequence 199, App	897	4	13.3	130	10	US-09-815-242-10326	Sequence 10326, A
825	4	13.3	120	9	US-10-121-050-434	Sequence 434, App	898	4	13.3	130	10	US-09-815-242-11240	Sequence 11240, A
826	4	13.3	120	9	US-10-141-755-434	Sequence 434, App	899	4	13.3	130	10	US-09-815-242-11732	Sequence 13732, A
827	4	13.3	120	9	US-09-989-729A-199	Sequence 199, App	900	4	13.3	130	12	US-10-072-159-8	Sequence 8, Appli
828	4	13.3	120	9	US-09-990-440-199	Sequence 199, App	901	4	13.3	131	9	US-09-738-626-6745	Sequence 5745, Ap
829	4	13.3	120	9	US-09-991-654-199	Sequence 199, App	902	4	13.3	131	9	US-09-924-340-64	Sequence 64, Appli
830	4	13.3	120	9	US-09-997-349-199	Sequence 199, App	903	4	13.3	131	9	US-09-992-600A-64	Sequence 64, Appli
831	4	13.3	120	9	US-09-997-440-199	Sequence 199, App	904	4	13.3	131	9	US-10-125-540-532	Sequence 532, App
832	4	13.3	120	9	US-09-997-628-199	Sequence 199, App	905	4	13.3	131	10	US-09-764-864-532	Sequence 532, App
833	4	13.3	120	9	US-09-997-683-199	Sequence 199, App	906	4	13.3	131	10	US-09-925-301-1379	Sequence 1379, Ap
834	4	13.3	120	9	US-10-143-032-434	Sequence 434, App	907	4	13.3	131	10	US-09-825-414-58	Sequence 1379, Ap
835	4	13.3	120	10	US-09-205-658-215	Sequence 215, App	908	4	13.3	132	10	US-09-754-862-7	Sequence 58, Appli
836	4	13.3	120	10	US-09-801-115-8	Sequence 8, Appli	909	4	13.3	132	10	US-09-754-862-11	Sequence 11, Appli
837	4	13.3	120	10	US-09-989-722-199	Sequence 199, App	910	4	13.3	132	10	US-09-829-482-16	Sequence 16, Appli
838	4	13.3	120	10	US-09-989-723-199	Sequence 199, App	911	4	13.3	132	10	US-09-864-761-44228	Sequence 44228, A
839	4	13.3	120	10	US-09-989-279-199	Sequence 199, App	912	4	13.3	132	10	US-09-925-301-963	Sequence 963, App
840	4	13.3	120	10	US-09-989-727-199	Sequence 199, App	913	4	13.3	132	12	US-10-062-254-40	Sequence 40, Appli
841	4	13.3	120	10	US-09-989-731-199	Sequence 199, App	914	4	13.3	133	9	US-09-866-050A-712	Sequence 712, App
842	4	13.3	120	10	US-09-989-732-199	Sequence 199, App	915	4	13.3	133	12	US-10-062-994-20	Sequence 20, Appli
843	4	13.3	120	10	US-09-991-073-199	Sequence 199, App	916	4	13.3	133	12	US-10-062-624-7	Sequence 7, Appli
844	4	13.3	120	10	US-09-990-442-199	Sequence 199, App	917	4	13.3	133	12	US-10-062-994-20	Sequence 20, Appli
845	4	13.3	120	10	US-09-991-163-199	Sequence 199, App	918	4	13.3	134	9	US-10-079-623-164	Sequence 364, App
846	4	13.3	120	10	US-09-990-456-199	Sequence 199, App	919	4	13.3	135	10	US-09-864-761-47958	Sequence 47958, A
847	4	13.3	120	10	US-09-990-456-199	Sequence 199, App	920	4	13.3	135	10	US-09-908-322-51	Sequence 51, Appli
848	4	13.3	120	10	US-09-989-721-199	Sequence 199, App	921	4	13.3	135	10	US-09-964-864-983	Sequence 983, App
849	4	13.3	121	9	US-09-989-721-199	Sequence 199, App	922	4	13.3	135	10	US-09-949-192-21	Sequence 21, Appli
850	4	13.3	121	9	US-10-028-072-48	Sequence 48, Appli	923	4	13.3	136	9	US-09-764-668-710	Sequence 710, App
851	4	13.3	121	9	US-10-121-049-48	Sequence 48, Appli	924	4	13.3	136	9	US-09-965-828-24	Sequence 24, Appli
852	4	13.3	121	9	US-10-123-904-48	Sequence 48, Appli	925	4	13.3	136	9	US-09-820-843A-66	Sequence 66, Appli
853	4	13.3	121	9	US-10-140-470-48	Sequence 48, Appli	926	4	13.3	136	9	US-10-235-674-13	Sequence 13, Appli
854	4	13.3	121	9	US-10-175-746-48	Sequence 48, Appli	927	4	13.3	136	9	US-10-101-664A-763	Sequence 763, App
855	4	13.3	121	9	US-10-176-918-48	Sequence 48, Appli	928	4	13.3	136	9	US-10-091-504-1064	Sequence 1064, Ap
856	4	13.3	121	9	US-10-176-921-48	Sequence 48, Appli	929	4	13.3	136	10	US-09-728-479-9	Sequence 9, Appli
857	4	13.3	121	9	US-10-137-865-48	Sequence 48, Appli	930	4	13.3	136	10	US-09-864-761-33497	Sequence 33497, A
858	4	13.3	121	9	US-10-140-474-48	Sequence 48, Appli	931	4	13.3	136	10	US-09-864-761-36438	Sequence 36438, A
859	4	13.3	121	9	US-10-142-431-48	Sequence 48, Appli	932	4	13.3	136	10	US-09-864-761-38438	Sequence 38438, A
860	4	13.3	121	9	US-10-143-114-48	Sequence 48, Appli	933	4	13.3	136	10	US-09-814-122-67	Sequence 67, Appli
861	4	13.3	121	9	US-10-140-002-48	Sequence 48, Appli	934	4	13.3	136	10	US-09-764-869-1064	Sequence 1064, Ap
862	4	13.3	121	9	US-10-142-419-48	Sequence 48, Appli	935	4	13.3	136	10	US-09-867-550-396	Sequence 396, App
863	4	13.3	121	9	US-10-123-362-48	Sequence 48, Appli	936	4	13.3	136	10	US-09-263-689-13	Sequence 13, Appli
864	4	13.3	121	9	US-10-142-423-48	Sequence 48, Appli	937	4	13.3	136	10	US-09-948-018-6	Sequence 6, Appli
865	4	13.3	121	9	US-10-121-050-48	Sequence 48, Appli	938	4	13.3	137	10	US-09-864-761-44668	Sequence 44668, A
866	4	13.3	121	9	US-10-141-755-48	Sequence 48, Appli	939	4	13.3	137	10	US-09-815-242-5865	Sequence 5865, Ap
867	4	13.3	121	9	US-10-091-504-899	Sequence 899, App	940	4	13.3	137	10	US-09-867-550-832	Sequence 832, App
868	4	13.3	121	9	US-10-143-032-48	Sequence 48, Appli	941	4	13.3	138	9	US-10-013-379-14	Sequence 14, Appli
869	4	13.3	121	10	US-09-764-869-899	Sequence 899, App	942	4	13.3	138	9	US-09-738-626-6238	Sequence 6238, Ap
870	4	13.3	122	10	US-09-854-799-30	Sequence 30, Appli	943	4	13.3	139	10	US-09-815-242-13195	Sequence 13195, A
871	4	13.3	122	10	US-09-775-932-10	Sequence 10, Appli	944	4	13.3	139	10	US-09-752-820A-5	Sequence 5, Appli
872	4	13.3	122	9	US-09-734-017A-56	Sequence 56, Appli	945	4	13.3	139	10	US-09-752-820A-7	Sequence 7, Appli
873	4	13.3	123	9	US-09-866-050A-292	Sequence 292, App	946	4	13.3	139	10	US-09-813-319A-5	Sequence 5, Appli
874	4	13.3	123	10	US-09-864-761-33505	Sequence 33505, A	947	4	13.3	139	10	US-09-813-319A-7	Sequence 7, Appli
875	4	13.3	124	10	US-09-864-761-40032	Sequence 40032, A	948	4	13.3	140	9	US-10-080-980-5	Sequence 5, Appli
876	4	13.3	124	10	US-09-815-248-8	Sequence 8, Appli	949	4	13.3	140	10	US-09-867-550-1252	Sequence 1252, Ap
877	4	13.3	125	9	US-10-102-806-491	Sequence 491, App	950	4	13.3	141	10	US-09-972-086-9	Sequence 9, Appli
878	4	13.3	125	10	US-09-854-799-54	Sequence 54, App	951	4	13.3	141	10	US-09-963-791-14	Sequence 14, Appli
879	4	13.3	126	10	US-09-887-855-6	Sequence 6, Appli	952	4	13.3	142	10	US-09-940-497-4	Sequence 4, Appli
880	4	13.3	126	10	US-09-887-855-6	Sequence 6, Appli	953	4	13.3	143	9	US-10-063-487-342	Sequence 342, App
881	4	13.3	126	10	US-09-815-248-6	Sequence 6, Appli	954	4	13.3	143	10	US-09-767-041-42	Sequence 42, Appli
882	4	13.3	126	10	US-09-949-972-4	Sequence 4, Appli	955	4	13.3	144	9	US-09-895-913A-156	Sequence 156, App
883	4	13.3	126	12	US-10-027-450-44	Sequence 44, Appli	956	4	13.3	144	12	US-10-062-254-34	Sequence 34, Appli
884	4	13.3	127	9	US-09-738-626-4452	Sequence 4452, Ap	957	4	13.3	145	9	US-10-101-654A-106	Sequence 106, App
885	4	13.3	127	10	US-09-867-550-762	Sequence 762, App	958	4	13.3	145	9	US-10-092-154A-555	Sequence 555, App
886	4	13.3	127	10	US-09-865-159-10	Sequence 10, Appli	959	4	13.3	145	10	US-09-764-864-137	Sequence 137, Ap
887	4	13.3	128	9	US-10-013-379-32	Sequence 32, Appli	960	4	13.3	145	10	US-09-764-864-137	Sequence 137, Ap
888	4	13.3	128	9	US-09-809-391-39	Sequence 39, Appli	961	4	13.3	146	9	US-09-978-895A-221	Sequence 221, App
889	4	13.3	128	10	US-09-741-669-429	Sequence 429, App	962	4	13.3	146	9	US-09-978-897-221	Sequence 221, App
890	4	13.3	128	10	US-09-747-835A-9	Sequence 9, Appli	963	4	13.3	146	9	US-09-978-192A-221	Sequence 221, App
891	4	13.3	129	9	US-09-854-133-214	Sequence 214, App	964	4	13.3	146	9	US-09-999-832A-221	Sequence 221, App
892	4	13.3	129	9	US-09-854-133-214	Sequence 214, App	965	4	13.3	146	9	US-09-978-189-221	Sequence 221, App
893	4	13.3	129	10	US-09-864-761-42770	Sequence 42770, A	966	4	13.3	146	9	US-10-028-072-408	Sequence 408, App
894	4	13.3	129	10	US-09-815-242-11228	Sequence 11228, A	967	4	13.3	146	9	US-10-121-049-408	Sequence 408, App
895	4	13.3	129	10	US-09-738-973-214	Sequence 214, App	968	4	13.3	146	9	US-10-123-904-408	Sequence 408, App

969 4 13.3 146 9 US-10-140-470-408 Sequence 408, App  
970 4 13.3 146 9 US-10-175-746-408 Sequence 408, App  
971 4 13.3 146 9 US-10-176-918-408 Sequence 408, App  
972 4 13.3 146 9 US-10-176-921-408 Sequence 408, App  
973 4 13.3 146 9 US-10-177-865-408 Sequence 408, App  
974 4 13.3 146 9 US-10-140-474-408 Sequence 408, App  
975 4 13.3 146 9 US-10-142-431-408 Sequence 408, App  
976 4 13.3 146 9 US-10-143-114-408 Sequence 408, App  
977 4 13.3 146 9 US-10-140-002-408 Sequence 408, App  
978 4 13.3 146 9 US-09-978-608A-221 Sequence 221, App  
979 4 13.3 146 9 US-10-142-419-408 Sequence 408, App  
980 4 13.3 146 9 US-09-978-191A-221 Sequence 221, App  
981 4 13.3 146 9 US-09-978-403A-221 Sequence 221, App  
982 4 13.3 146 9 US-09-978-564A-221 Sequence 221, App  
983 4 13.3 146 9 US-09-978-585A-221 Sequence 221, App  
984 4 13.3 146 9 US-10-017-081A-221 Sequence 221, App  
985 4 13.3 146 9 US-10-123-262-408 Sequence 408, App  
986 4 13.3 146 9 US-10-142-423-408 Sequence 408, App  
987 4 13.3 146 9 US-09-978-824-221 Sequence 221, App  
988 4 13.3 146 9 US-09-981-915A-221 Sequence 221, App  
989 4 13.3 146 9 US-09-999-833A-221 Sequence 221, App  
990 4 13.3 146 9 US-10-097-065-313 Sequence 313, App  
991 4 13.3 146 9 US-10-121-050-408 Sequence 408, App  
992 4 13.3 146 9 US-10-141-755-408 Sequence 408, App  
993 4 13.3 146 9 US-10-167-749-221 Sequence 221, App  
994 4 13.3 146 9 US-09-918-585A-221 Sequence 221, App  
995 4 13.3 146 9 US-10-143-032-408 Sequence 408, App  
996 4 13.3 146 10 US-09-765-205-42 Sequence 42, App  
997 4 13.3 146 10 US-09-815-242-11937 Sequence 11937, A  
998 4 13.3 147 9 US-10-043-487-325 Sequence 325, App  
999 4 13.3 147 10 US-09-811-284-172 Sequence 172, App  
1000 4 13.3 148 9 US-09-736-457-338 Sequence 338, App

## ALIGNMENTS

RESULT 1  
US-09-987-190-2  
Sequence 2, Application US/09987190  
Patent No. US20020058293A1  
GENERAL INFORMATION:  
APPLICANT: TAKESAKO, Kazutosh  
APPLICANT: MIZUTANI, Shigetoshi  
APPLICANT: ENDO, Masahiro  
APPLICANT: KATO, Ikunoshin  
TITLE OF INVENTION: FUNGAL ANTIGENS AND PROCESS FOR PRODUCING THE SAME  
FILE REFERENCE: 1422-0502P  
CURRENT APPLICATION NUMBER: US/09/987,190  
CURRENT FILING DATE: 2001-11-13  
PRIOR APPLICATION NUMBER: 09/262,856  
PRIOR FILING DATE: 1999-03-04  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 2  
LENGTH: 30  
TYPE: PRT  
ORGANISM: Candida albicans  
FEATURE:  
NAME/KEY: Unsure  
LOCATION: (1)..(30)  
OTHER INFORMATION: any Xaa = any amino acid, unknown or other  
US-09-987-190-2

Query Match 90.0%; Score 27; DB 10; Length 30;  
Best Local Similarity 100.0%; Pred. No. 3.4e-22;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 KYSLPELDYFSATSEPTYSQINRIXT 28  
Db 1 KYSLPELDYFSATSEPTYSQINRIXT 28

RESULT 2  
US-09-864-761-43011  
Sequence 43011, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aecmca-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 43011  
LENGTH: 50  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC007510.4  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.76  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.61  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.7  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.76  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.72  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.58  
OTHER INFORMATION: EST HUMAN HIT: BF667240.1, EVALUATE 1.30e+00  
OTHER INFORMATION: SWISSPROT HIT: P87145, EVALUATE 6.00e+00  
US-09-864-761-43011

Query Match 20.0%; Score 6; DB 10; Length 50;  
Best Local Similarity 100.0%; Pred. No. 8.7; Indels 0; Gaps 0;  
Matches 6; Conservative 0; Mismatches 0;

Cy 12 SATPEY 17

Db 10 SATPEY 15

RESULT 3

US-10-078-929-24

Sequence 24, Application US/10078929  
Patent No. US20020152497A1  
GENERAL INFORMATION:  
APPLICANT: Rafalski, Antoni  
APPLICANT: Miao, Guo-Hua  
APPLICANT: Sakai, Hajime  
APPLICANT: Falcio, Saverio Carl  
APPLICANT: Odeli, Joan T.  
APPLICANT: Meyers, Blake  
APPLICANT: Thorpe, Catherine  
APPLICANT: Meng, Zude  
TITLE OF INVENTION: Nucleic Acid Fragments Encoding Proteins Involved in  
TITLE OF INVENTION: Stress Response  
FILE REFERENCE: B1357 US NA  
CURRENT APPLICATION NUMBER: US/10/078,929  
CURRENT FILING DATE: 2002-02-19

PRIOR APPLICATION NUMBER: 09/566,394  
PRIOR FILING DATE: 2000-05-05  
PRIOR APPLICATION NUMBER: 60/133038  
PRIOR FILING DATE: 1999-05-07  
PRIOR APPLICATION NUMBER: 60/133042  
PRIOR FILING DATE: 1999-05-07  
PRIOR APPLICATION NUMBER: 60/133427  
PRIOR FILING DATE: 1999-05-11  
PRIOR APPLICATION NUMBER: 60/133437  
PRIOR FILING DATE: 1999-05-11  
PRIOR APPLICATION NUMBER: 60/133428  
PRIOR FILING DATE: 1999-05-11  
PRIOR APPLICATION NUMBER: 60/133438  
PRIOR FILING DATE: 1999-05-11  
PRIOR APPLICATION NUMBER: 60/133436  
PRIOR FILING DATE: 1999-05-11  
PRIOR APPLICATION NUMBER: 60/133667  
PRIOR FILING DATE: 1999-06-04  
NUMBER OF SEQ ID NOS: 208  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 24  
LENGTH: 65  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: (60)  
NAME/KEY: UNSURE  
LOCATION: (63)  
US-10-078-929-24

Query Match 20.0%; Score 6; DB 12; Length 65;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PELDYR 10  
|||||  
Db 44 PELDYR 49

RESULT 4

US-09-981-876-166  
Sequence 166, Application US/09981876  
Patent No. US2002016469A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 70 Human Secreted Proteins  
FILE REFERENCE: P2001P1  
CURRENT APPLICATION NUMBER: US/09/981,876  
CURRENT FILING DATE: 2001-10-19

PRIOR APPLICATION NUMBER: 09/148,545  
PRIOR FILING DATE: 1998-09-04  
PRIOR APPLICATION NUMBER: 60/040,162  
PRIOR FILING DATE: 1997-03-07  
PRIOR APPLICATION NUMBER: 60/040,333  
PRIOR FILING DATE: 1997-03-07  
PRIOR APPLICATION NUMBER: 60/038,621  
PRIOR FILING DATE: 1997-03-07  
PRIOR APPLICATION NUMBER: 60/040,161  
PRIOR FILING DATE: 1997-03-07  
PRIOR APPLICATION NUMBER: 60/040,626  
PRIOR FILING DATE: 1997-03-07  
PRIOR APPLICATION NUMBER: 60/040,334  
PRIOR FILING DATE: 1997-03-07  
PRIOR APPLICATION NUMBER: 60/040,336  
PRIOR FILING DATE: 1997-03-07  
PRIOR APPLICATION NUMBER: 60/040,163  
PRIOR FILING DATE: 1997-03-07  
PRIOR APPLICATION NUMBER: 60/047,615  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,600  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,597  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,502  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,633  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,583  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,617  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,618  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,503  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,592  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,581  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,584  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,500  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,587  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,492  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,598  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,613  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,582  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,596  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,612  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,632  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,601  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/043,580  
PRIOR FILING DATE: 1997-04-11  
PRIOR APPLICATION NUMBER: 60/043,568  
PRIOR FILING DATE: 1997-04-11  
PRIOR APPLICATION NUMBER: 60/043,314  
PRIOR FILING DATE: 1997-04-11  
PRIOR APPLICATION NUMBER: 60/043,569  
PRIOR FILING DATE: 1997-04-11  
PRIOR APPLICATION NUMBER: 60/043,311  
PRIOR FILING DATE: 1997-04-11  
PRIOR APPLICATION NUMBER: 60/043,671

;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,674  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,669  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,312  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,313  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,672  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,315  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/048,974  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/056,886  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,877  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,889  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,893  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,630  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,878  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,662  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,872  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,882  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,637  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,903  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,888  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,879  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,880  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,894  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,911  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,636  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,874  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,910  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,864  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,631  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,845  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,892  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/047,595  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/057,761  
;; PRIOR FILING DATE: 05-Sep-1997  
;; PRIOR APPLICATION NUMBER: 60/047,599  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,588  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,585  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,586  
;; PRIOR FILING DATE: 1997-05-23

;; PRIOR APPLICATION NUMBER: 60/047,590  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,594  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,589  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,593  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,614  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/043,578  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,576  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/047,501  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/043,670  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/056,632  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,664  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,876  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,881  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,909  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,875  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,862  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,887  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,908  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/048,964  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/057,650  
;; PRIOR FILING DATE: 1997-09-05  
;; PRIOR APPLICATION NUMBER: 60/056,884  
;; PRIOR FILING DATE: 1997-08-22  
;; NUMBER OF SEQ ID NOS: 280  
;; SOFTWARE: Patentin Ver. 2.0  
;; SEQ ID NO 166  
;; LENGTH: 81

Query Match 20.0%; Score 6; DB 9; Length 81;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 ATEPYI 18  
Db 71 ATEPYI 76

RESULT 5  
US-09-148-545-166  
; Sequence 166, Application US/09148545  
; Publication No. US20030027132A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: 70 Human Secreted Proteins  
; FILE REFERENCE: P2001PI  
; CURRENT FILING DATE: US/09/148,545  
; EARLIER APPLICATION NUMBER: PCT/US98/04482  
; EARLIER FILING DATE: 1998-03-06  
; EARLIER APPLICATION NUMBER: 60/040,162  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/040,333  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/038,621

EARLIER	FILING DATE:	1997-03-07
EARLIER	APPLICATION NUMBER:	60/040,161
EARLIER	FILING DATE:	1997-03-07
EARLIER	APPLICATION NUMBER:	60/040,626
EARLIER	FILING DATE:	1997-03-07
EARLIER	APPLICATION NUMBER:	60/040,334
EARLIER	FILING DATE:	1997-03-07
EARLIER	APPLICATION NUMBER:	60/040,336
EARLIER	FILING DATE:	1997-03-07
EARLIER	APPLICATION NUMBER:	60/040,163
EARLIER	FILING DATE:	1997-03-07
EARLIER	APPLICATION NUMBER:	60/047,615
EARLIER	FILING DATE:	1997-05-23
EARLIER	APPLICATION NUMBER:	60/047,600
EARLIER	FILING DATE:	1997-05-23
EARLIER	APPLICATION NUMBER:	60/047,597
EARLIER	FILING DATE:	1997-05-23
EARLIER	APPLICATION NUMBER:	60/047,583
EARLIER	FILING DATE:	1997-05-23
EARLIER	APPLICATION NUMBER:	60/047,617
EARLIER	FILING DATE:	1997-05-23
EARLIER	APPLICATION NUMBER:	60/047,618
EARLIER	FILING DATE:	1997-05-23
EARLIER	APPLICATION NUMBER:	60/047,503
EARLIER	FILING DATE:	1997-05-23
EARLIER	APPLICATION NUMBER:	60/047,592
EARLIER	FILING DATE:	1997-05-23
EARLIER	APPLICATION NUMBER:	60/047,581
EARLIER	FILING DATE:	1997-05-23
EARLIER	APPLICATION NUMBER:	60/047,584
EARLIER	FILING DATE:	1997-05-23
EARLIER	APPLICATION NUMBER:	60/047,500
EARLIER	FILING DATE:	1997-05-23
EARLIER	APPLICATION NUMBER:	60/047,587
EARLIER	FILING DATE:	1997-05-23
EARLIER	APPLICATION NUMBER:	60/047,492
EARLIER	FILING DATE:	1997-05-23
EARLIER	APPLICATION NUMBER:	60/047,598
EARLIER	FILING DATE:	1997-05-23
EARLIER	APPLICATION NUMBER:	60/047,613
EARLIER	FILING DATE:	1997-05-23
EARLIER	APPLICATION NUMBER:	60/047,582
EARLIER	FILING DATE:	1997-05-23
EARLIER	APPLICATION NUMBER:	60/047,596
EARLIER	FILING DATE:	1997-05-23
EARLIER	APPLICATION NUMBER:	60/047,612
EARLIER	FILING DATE:	1997-05-23
EARLIER	APPLICATION NUMBER:	60/047,632
EARLIER	FILING DATE:	1997-05-23
EARLIER	APPLICATION NUMBER:	60/047,601
EARLIER	FILING DATE:	1997-05-23
EARLIER	APPLICATION NUMBER:	60/043,580
EARLIER	FILING DATE:	1997-04-11
EARLIER	APPLICATION NUMBER:	60/043,568
EARLIER	FILING DATE:	1997-04-11
EARLIER	APPLICATION NUMBER:	60/043,314
EARLIER	FILING DATE:	1997-04-11
EARLIER	APPLICATION NUMBER:	60/043,569
EARLIER	FILING DATE:	1997-04-11
EARLIER	APPLICATION NUMBER:	60/043,311
EARLIER	FILING DATE:	1997-04-11
EARLIER	APPLICATION NUMBER:	60/043,671
EARLIER	FILING DATE:	1997-04-11
EARLIER	APPLICATION NUMBER:	60/043,674
EARLIER	FILING DATE:	1997-04-11
EARLIER	APPLICATION NUMBER:	60/043,669
EARLIER	FILING DATE:	1997-04-11
EARLIER	APPLICATION NUMBER:	60/043,312
EARLIER	FILING DATE:	1997-04-11

1	EARLIER	APPLICATION NUMBER: 60/043, 3113
2	EARLIER	FILING DATE: 1997-04-11
3	EARLIER	APPLICATION NUMBER: 60/043, 6722
4	EARLIER	FILING DATE: 1997-04-11
5	EARLIER	APPLICATION NUMBER: 60/043, 315
6	EARLIER	FILING DATE: 1997-04-11
7	EARLIER	APPLICATION NUMBER: 60/048, 974
8	EARLIER	FILING DATE: 1997-06-06
9	EARLIER	APPLICATION NUMBER: 60/056, 8866
10	EARLIER	FILING DATE: 1997-08-22
11	EARLIER	APPLICATION NUMBER: 60/056, 8777
12	EARLIER	FILING DATE: 1997-08-22
13	EARLIER	APPLICATION NUMBER: 60/056, 8899
14	EARLIER	FILING DATE: 1997-08-22
15	EARLIER	APPLICATION NUMBER: 60/056, 6522
16	EARLIER	FILING DATE: 1997-08-22
17	EARLIER	APPLICATION NUMBER: 60/056, 8882
18	EARLIER	FILING DATE: 1997-08-22
19	EARLIER	APPLICATION NUMBER: 60/056, 6377
20	EARLIER	FILING DATE: 1997-08-22
21	EARLIER	APPLICATION NUMBER: 60/056, 9030
22	EARLIER	FILING DATE: 1997-08-22
23	EARLIER	APPLICATION NUMBER: 60/056, 8888
24	EARLIER	FILING DATE: 1997-08-22
25	EARLIER	APPLICATION NUMBER: 60/056, 8799
26	EARLIER	FILING DATE: 1997-08-22
27	EARLIER	APPLICATION NUMBER: 60/056, 8800
28	EARLIER	FILING DATE: 1997-08-22
29	EARLIER	APPLICATION NUMBER: 60/056, 8944
30	EARLIER	FILING DATE: 1997-08-22
31	EARLIER	APPLICATION NUMBER: 60/056, 8744
32	EARLIER	FILING DATE: 1997-08-22
33	EARLIER	APPLICATION NUMBER: 60/056, 9111
34	EARLIER	FILING DATE: 1997-08-22
35	EARLIER	APPLICATION NUMBER: 60/056, 6366
36	EARLIER	FILING DATE: 1997-08-22
37	EARLIER	APPLICATION NUMBER: 60/056, 8744
38	EARLIER	FILING DATE: 1997-08-22
39	EARLIER	APPLICATION NUMBER: 60/056, 9100
40	EARLIER	FILING DATE: 1997-08-22
41	EARLIER	APPLICATION NUMBER: 60/056, 8644
42	EARLIER	FILING DATE: 1997-08-22
43	EARLIER	APPLICATION NUMBER: 60/056, 6311
44	EARLIER	FILING DATE: 1997-08-22
45	EARLIER	APPLICATION NUMBER: 60/056, 8455
46	EARLIER	FILING DATE: 1997-08-22
47	EARLIER	APPLICATION NUMBER: 60/056, 8922
48	EARLIER	FILING DATE: 1997-08-22
49	EARLIER	APPLICATION NUMBER: 60/047, 5595
50	EARLIER	FILING DATE: 1997-05-23
51	EARLIER	APPLICATION NUMBER: 60/047, 7611
52	EARLIER	FILING DATE: 05-Sep-1997
53	EARLIER	APPLICATION NUMBER: 60/047, 5999
54	EARLIER	FILING DATE: 1997-05-23
55	EARLIER	APPLICATION NUMBER: 60/047, 5888
56	EARLIER	FILING DATE: 1997-05-23
57	EARLIER	APPLICATION NUMBER: 60/047, 5855
58	EARLIER	FILING DATE: 1997-05-23
59	EARLIER	APPLICATION NUMBER: 60/047, 5866
60	EARLIER	FILING DATE: 1997-05-23
61	EARLIER	APPLICATION NUMBER: 60/047, 5900
62	EARLIER	FILING DATE: 1997-05-23
63	EARLIER	APPLICATION NUMBER: 60/047, 5944
64	EARLIER	FILING DATE: 1997-05-23
65	EARLIER	APPLICATION NUMBER: 60/047, 5899
66	EARLIER	FILING DATE: 1997-05-23
67	EARLIER	APPLICATION NUMBER: 60/047, 5933

EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,614  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043,578  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,576  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/047,501  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043,670  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/056,632  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,664  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,876  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,881  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,909  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,875  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,862  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,887  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,908  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/048,964  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/057,650  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/056,884  
EARLIER FILING DATE: 1997-08-22  
NUMBER OF SEQ ID NOS: 280  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 166  
LENGTH: 81

Query Match 20.0%; Score 6; DB 9; Length 81;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 ATEPYI 18  
Db 71 ATEPYI 76

US-10-078-929-26  
Sequence 26, Application US/10078929  
Patent No. US20020152497A1  
GENERAL INFORMATION:  
APPLICANT: Rafalski, Antoni  
APPLICANT: Miao, Guo-Hua  
APPLICANT: Falco, Saverio Carl  
APPLICANT: Sakai, Hajime  
APPLICANT: Pamodu, Omolayo O.  
APPLICANT: Odell, Joan T.  
APPLICANT: Meyers, Blake  
APPLICANT: Thorpe, Catherine  
APPLICANT: Meng, Zude  
TITLE OF INVENTION: Nucleic Acid Fragments Encoding Proteins Involved in  
FILE REFERENCE: BB1357 US NA  
CURRENT APPLICATION NUMBER: US/10/078,929  
CURRENT FILING DATE: 2002-02-19  
PRIOR APPLICATION NUMBER: 09/566,394  
PRIOR FILING DATE: 2000-05-05  
PRIOR APPLICATION NUMBER: 60/133038  
PRIOR FILING DATE: 1999-05-07  
PRIOR APPLICATION NUMBER: 60/133042

PRIOR FILING DATE: 1999-05-07  
PRIOR APPLICATION NUMBER: 60/133427  
PRIOR FILING DATE: 1999-05-11  
PRIOR APPLICATION NUMBER: 60/133437  
PRIOR FILING DATE: 1999-05-11  
PRIOR APPLICATION NUMBER: 60/133428  
PRIOR FILING DATE: 1999-05-11  
PRIOR APPLICATION NUMBER: 60/133438  
PRIOR FILING DATE: 1999-05-11  
PRIOR APPLICATION NUMBER: 60/133436  
PRIOR FILING DATE: 1999-05-11  
PRIOR APPLICATION NUMBER: 60/137667  
PRIOR FILING DATE: 1999-06-04  
NUMBER OF SEQ ID NOS: 208  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 26  
LENGTH: 98  
TYPE: PRT  
ORGANISM: Glycine max  
US-10-078-929-26

Query Match 20.0%; Score 6; DB 12; Length 98;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 PELDYE 10  
Db 64 PELDYE 69

US-09-840-787-49  
Sequence 49, Application US/09840787  
Patent No. US20020058264A1  
GENERAL INFORMATION:  
APPLICANT: Lal, Preeti  
Hillman, Jennifer L.  
Bandman, Olga  
Shah, Purvi  
Au-Young, Janice  
Yue, Henry  
Guegler, Karl J.  
Corley, Neil C.  
TITLE OF INVENTION: HUMAN REGULATORY MOLECULES  
NUMBER OF SEQUENCES: 98  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/840,787  
FILING DATE: 23-Apr-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/518,865  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0356 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 49:



## SEQUENCE CHARACTERISTICS:

LENGTH: 137 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: SPINNOUT10  
CLONE: 3339274  
SEQUENCE DESCRIPTION: SEQ ID NO: 49 :  
US-09-840-787-49

## Query Match

Best Local Similarity 20.0%; Score 6; DB 10; Length 137;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 YISGOI 22

Db 32 YISGOI 37

## RESULT 8

US-09-925-300-975  
Sequence 975, Application US/09925300  
Patent No. US20020151681A1  
GENERAL INFORMATION:  
APPLICANT: Craig Rosen,  
APPLICANT: Steve Ruben,  
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
FILE REFERENCE: PA101  
CURRENT APPLICATION NUMBER: US/09/925,300  
CURRENT FILING DATE: 2001-08-10  
PRIOR APPLICATION NUMBER: PCT/US00/05988  
PRIOR FILING DATE: 2000-03-08  
PRIOR APPLICATION NUMBER: 60/124,270  
PRIOR FILING DATE: 1999-03-12  
NUMBER OF SEQ ID NOS: 1890  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 975  
LENGTH: 146  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-925-300-975

## Query Match

Best Local Similarity 20.0%; Score 6; DB 10; Length 146;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 YISGOI 22

Db 41 YISGOI 46

## RESULT 9

US-09-987-190-6  
Sequence 6, Application US/09987190  
Patent No. US20020058293A1  
GENERAL INFORMATION:  
APPLICANT: TAKESAKO, Kazutoh  
APPLICANT: MIZUTANI, Shigetoshi  
APPLICANT: ENDO, Masahito  
APPLICANT: KADO, Ikunoshin  
TITLE OF INVENTION: FUNGAL ANTIGENS AND PROCESS FOR PRODUCING THE SAME  
FILE REFERENCE: 1422-0502P  
CURRENT APPLICATION NUMBER: US/09/987,190  
CURRENT FILING DATE: 2001-11-13  
PRIOR APPLICATION NUMBER: 09/262,856  
PRIOR FILING DATE: 1999-03-04  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 6  
LENGTH: 188  
TYPE: PRT  
ORGANISM: Candida albicans

US-09-987-190-6

Query Match 20.0%; Score 6; DB 10; Length 188;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 ATEPYI 18

Db 1 ATEPYI 6

## RESULT 10

US-09-738-626-6727  
Sequence 6727, Application US/09738626  
Publication No. US20020197605A1  
GENERAL INFORMATION:  
APPLICANT: NAKAGAWA, SATOSHI  
APPLICANT: MIZOGUCHI, HIROSHI  
APPLICANT: ANDO, SEIKO  
APPLICANT: HAYASHI, MIKIRO  
APPLICANT: OCHIAI, KEIKO  
APPLICANT: YOKOI, HARUHIKO  
APPLICANT: TATEISHI, NAOKO  
APPLICANT: SENOH, AKIHIRO  
APPLICANT: IKEDA, MASATO  
APPLICANT: OZAKI, AKIO  
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
FILE REFERENCE: 249-125  
CURRENT APPLICATION NUMBER: US/09/738,626  
CURRENT FILING DATE: 2000-12-18  
PRIOR APPLICATION NUMBER: JP 99/377484  
PRIOR FILING DATE: 1999-12-16  
PRIOR APPLICATION NUMBER: JP 00/159162  
PRIOR FILING DATE: 2000-04-07  
PRIOR APPLICATION NUMBER: JP 00/280988  
PRIOR FILING DATE: 2000-08-03  
NUMBER OF SEQ ID NOS: 7059  
SOFTWARE: PatentIn ver. 3.0  
SEQ ID NO 6727  
LENGTH: 200  
TYPE: PRT  
ORGANISM: Corynebacterium glutamicum  
US-09-738-626-6727

## Query Match

Best Local Similarity 20.0%; Score 6; DB 9; Length 200;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 LPELDY 9

Db 6 LPELDY 11

## RESULT 11

US-09-818-564-2  
Sequence 2, Application US/09818564  
Patent No. US20020137151A1  
GENERAL INFORMATION:  
APPLICANT: MERRIAM, MORIEL  
APPLICANT: GUYONVARCH, ARNEL  
APPLICANT: MARX, ACHIM  
TITLE OF INVENTION: A PROCESS FOR THE FERMENTATIVE PREPARATION OF METABOLIC PRODUCTS AND FOR THE NUCLEOTIDE SEQUENCES ENCODING FOR  
FILE REFERENCE: 21123/278411/MAS  
CURRENT APPLICATION NUMBER: US/09/818,564  
CURRENT FILING DATE: 2001-03-28  
PRIOR APPLICATION NUMBER: 09/373,731  
PRIOR FILING DATE: 1999-08-13  
NUMBER OF SEQ ID NOS: 8  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 200

TYPE: PRT  
ORGANISM: Corynebacterium melassecola  
OTHER INFORMATION: ATCC 17965  
US-09-818-564-2

Query Match 20.0%; Score 6; DB 10; Length 200;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 LPELDY 9  
Db 6 LPELDY 11

RESULT 12  
US-09-909-567B-47  
Sequence 47, Application US/099095567B  
Publication No. US20030022257A1  
GENERAL INFORMATION:  
APPLICANT: Macina, Roberto A.  
APPLICANT: Nair, Manoj  
APPLICANT: Chen, Seiyu

TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes

FILE REFERENCE: DEX-0214  
CURRENT APPLICATION NUMBER: US/09/909,567B  
CURRENT FILING DATE: 2001-07-20  
PRIOR APPLICATION NUMBER: 60/219,834  
PRIOR FILING DATE: 2000-07-21  
NUMBER OF SEQ ID NOS: 56  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 47

LENGTH: 1596  
TYPE: PRT

ORGANISM: Homo sapien  
US-09-909-567B-47

Query Match 20.0%; Score 6; DB 9; Length 1596;  
Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 SLEPDL 8  
Db 1356 SLEPDL 1361

RESULT 13  
US-09-813-153-211

Sequence 211, Application US/09813153  
Publication No. US20030045459A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.

TITLE OF INVENTION: 67 Human secreted proteins

FILE REFERENCE: P2023

CURRENT APPLICATION NUMBER: US/09/813,153

CURRENT FILING DATE: 2001-03-21

PRIOR APPLICATION NUMBER: US/09/363,044

PRIOR FILING DATE: 1999-07-29

PRIOR APPLICATION NUMBER: 60/073,160

PRIOR FILING DATE: 1998-01-30

PRIOR APPLICATION NUMBER: 60/073,159

PRIOR FILING DATE: 1998-01-30

PRIOR APPLICATION NUMBER: 60/073,165

PRIOR FILING DATE: 1998-01-30

PRIOR APPLICATION NUMBER: 60/073,164

PRIOR FILING DATE: 1998-01-30

PRIOR APPLICATION NUMBER: 60/073,167

PRIOR FILING DATE: 1998-01-30

PRIOR APPLICATION NUMBER: 60/073,162

PRIOR FILING DATE: 1998-01-30

PRIOR APPLICATION NUMBER: 60/073,161

PRIOR FILING DATE: 1998-01-30

PRIOR APPLICATION NUMBER: 60/073,170

PRIOR FILING DATE: 1998-01-30

NUMBER OF SEQ ID NOS: 298  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 211  
LENGTH: 26  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-813-153-211

Query Match 16.7%; Score 5; DB 9; Length 26;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 SLEPL 7  
Db 13 SLEPL 17

RESULT 14  
US-09-864-761-48159  
Sequence 48159, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

FILE REFERENCE: Aeomica-x-1

CURRENT APPLICATION NUMBER: US/09/864,761

CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263,6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00666

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00667

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00669

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00665

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00668

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00663

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00662

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00661

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00670

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: US 60/234,687

PRIOR FILING DATE: 2000-09-21

PRIOR APPLICATION NUMBER: US 09/608,408

PRIOR FILING DATE: 2000-06-30

PRIOR APPLICATION NUMBER: US 09/774,203

PRIOR FILING DATE: 2001-01-29

NUMBER OF SEQ ID NOS: 49117

SOFTWARE: Anomax Sequence Listing Engine vers. 1.1

SEQ ID NO 48159

LENGTH: 33

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:  
OTHER INFORMATION: MAP TO AC009303.1  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.7  
OTHER INFORMATION: EST\_HUMAN HIT: A1902949.1, EVALU6 6.00e-11  
US-09-864-761-48159

Query Match 16.7%; Score 5; DB 10; Length 33;  
Best Local Similarity 100.0%; Pred. No. 68;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YSLPE 6  
Db 29 YSLPE 33

RESULT 15  
US-09-810-310-2  
Sequence 2, Application US/09810310  
Patent No. US2002004948A1  
GENERAL INFORMATION:  
APPLICANT: Khleif, Samir N.  
APPLICANT: Bezdek, Jay A.  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CO-STIMULATION OF  
FILE REFERENCE: 15280-415100US  
CURRENT FILING DATE: 2001-03-14  
PRIOR APPLICATION NUMBER: 60/189,396  
PRIOR FILING DATE: 2000-03-15  
NUMBER OF SEQ ID NOS: 61  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 39  
TYPE: PRT  
ORGANISM: Artificial Sequence  
OTHER INFORMATION: Description of Artificial Sequence: HIV PEPTIDE  
US-09-810-310-2

Query Match 16.7%; Score 5; DB 10; Length 39;  
Best Local Similarity 100.0%; Pred. No. 79;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 ISGQI 22  
Db 19 ISGQI 23

RESULT 16  
US-09-810-310-9  
Sequence 9, Application US/09810310  
Patent No. US2002004948A1  
GENERAL INFORMATION:  
APPLICANT: Khleif, Samir N.  
APPLICANT: Bezdek, Jay A.  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CO-STIMULATION OF  
FILE REFERENCE: 15280-415100US  
CURRENT FILING DATE: 2001-03-14  
PRIOR APPLICATION NUMBER: 60/189,396  
PRIOR FILING DATE: 2000-03-15  
NUMBER OF SEQ ID NOS: 61  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 9  
LENGTH: 39  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: HIV PEPTIDE  
US-09-810-310-9

Query Match 16.7%; Score 5; DB 10; Length 39;  
Best Local Similarity 100.0%; Pred. No. 79;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 ISGQI 22  
Db 19 ISGQI 23

RESULT 17  
US-09-864-761-34370  
Sequence 34370, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Hanzel, David R.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aecm1ca-X-1  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US/09/864,761  
PRIOR FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1  
SEQ ID NO 34370  
LENGTH: 39  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC007914.1  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3  
OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 1.1  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.6

OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 12  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.1  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.1  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2  
OTHER INFORMATION: EST\_HUMAN HIT: AV659020.1, EVALU6 6.00e-12  
US-09-864-761-34370

Query Match 16.7%; Score 5; DB 10; Length 39;  
Best Local Similarity 100.0%; Pred. No. 79;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SLEP 7  
Db 16 SLEP 20

RESULT 18  
US-09-764-904-63  
Sequence 63, Application US/09764904  
Patent No. US20020173454A1  
GENERAL INFORMATION:

APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PA122  
CURRENT APPLICATION NUMBER: US/09/764,904  
CURRENT FILING DATE: 2001-01-17  
Prior application data removed - consult PALM or file wrapper  
NUMBER OF SEQ ID NOS: 137  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 63  
LENGTH: 52  
TYPE: PRT

ORGANISM: Homo sapiens  
US-09-764-904-63

Query Match 16.7%; Score 5; DB 9; Length 52;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SLEP 7  
Db 8 SLEP 12

RESULT 19  
US-10-091-548-63

Sequence 63, Application US/10091548  
Publication No. US20030049703A1  
GENERAL INFORMATION:

APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PA122C1  
CURRENT APPLICATION NUMBER: US/10/091,548  
CURRENT FILING DATE: 2002-03-07  
NUMBER OF SEQ ID NOS: 137  
Prior application removed - See File Wrapper or Palm  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 63  
LENGTH: 52  
TYPE: PRT

ORGANISM: Homo sapiens  
US-10-091-548-63

Query Match 16.7%; Score 5; DB 9; Length 52;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SLEP 7  
Db 8 SLEP 12

RESULT 20  
US-09-764-860-581  
Sequence 581, Application US/09764860  
Patent No. US20020094953A1  
GENERAL INFORMATION:

APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PC008  
CURRENT APPLICATION NUMBER: US/09/764,860  
CURRENT FILING DATE: 2001-01-17  
Prior application data removed - consult PALM or file wrapper  
NUMBER OF SEQ ID NOS: 1198  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 581  
LENGTH: 52  
TYPE: PRT

ORGANISM: Homo sapiens  
US-09-764-860-581

Query Match 16.7%; Score 5; DB 10; Length 52;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SLEP 7  
Db 8 SLEP 12

RESULT 21  
US-09-864-761-34543  
Sequence 34543, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.  
APPLICANT: Hanzel, David K.  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aecmca-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263,6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687

```
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34543
; LENGTH: 65
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005089.2
; OTHER INFORMATION: EXPRESSED IN LONG, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.73
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.78
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
; OTHER INFORMATION: SWISSPROT HIT: Q84133, EVALU = 5.20e-01
; OTHER INFORMATION: EST_HUMAN HIT: BE258324.1, EVALU = 3.00e-32
; US-09-864-761-34543
```

```
Query Match          16.7%; Score 5; DB 10; Length 65;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 KYSLP 5
Db 49 KYSLP 53

RESULT 22
US-09-867-550-424
; Sequence 424, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehrtan, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells and
; FILE REFERENCE: Thereby
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 424
; LENGTH: 76
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-867-550-424
```

```
Query Match          16.7%; Score 5; DB 10; Length 76;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3 SLEPL 7
Db 65 SLEPL 69
```

```
RESULT 23
US-09-864-761-47808
; Sequence 47808, Application US/09864761
```

```
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 47808
; LENGTH: 77
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC011472.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.75
; OTHER INFORMATION: EST HUMAN HIT: BF36965.1, EVALU = 1.00e-05
; OTHER INFORMATION: SWISSPROT HIT: Q06639, EVALU = 2.50e+00
; US-09-864-761-47808
```

```
Query Match          16.7%; Score 5; DB 10; Length 77;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 18 ISGQI 22
Db 73 ISGQI 77
```

```
RESULT 24
US-09-764-887-257
; Sequence 257, Application US/09764887
; Patent No. US20020042096A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P413
; CURRENT APPLICATION NUMBER: US/09/764,887
; PRIOR FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 658
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 257
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (87)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-887-257

Query Match          16.7%; Score 5; DB 10; Length 88;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      12 SATPE 16
      |||||
Db      29 SATPE 33

RESULT 25
US-09-864-761-40079
; Sequence 40079, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
```

```

; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 40079
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004615.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
; OTHER INFORMATION: EST HUMAN HIT: AV118557.1, EVALUE 2.00e-19
; OTHER INFORMATION: SWISSPROT HIT: Q13591, EVALUE 2.00e-50
US-09-864-761-40079

Query Match          16.7%; Score 5; DB 10; Length 91;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 LPELD 8
      |||||
Db      76 LPELD 80

RESULT 26
US-10-101-464A-651
; Sequence 651, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; FILE REFERENCE: 11000.1020C2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 651
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-10-101-464A-651

Query Match          16.7%; Score 5; DB 9; Length 97;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 SLPEL 7
      |||||
Db      90 SLPEL 94

RESULT 27
US-09-764-877-1863
```

```

; Sequence 1863, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1863
; LENGTH: 105
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (47)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (68)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (95)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-877-1863

```

```

Query Match          16.7%; Score 5; DB 10; Length 105;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      3 SLPBL 7
        |||||
Db      14 SLPBL 18

```

```

RESULT 28
US-10-052-817-6
; Sequence 6, Application US/10052817
; Patent No. US20020114792A1
; GENERAL INFORMATION:
; APPLICANT: Tanzi, Rudolph E.
; APPLICANT: Kovacs, Dora
; APPLICANT: Saunders, Aleister J.
; TITLE OF INVENTION: Alpha-2-Macroglobulin Therapies and Drug Screening Methods for
; TITLE OF INVENTION: Alzheimer's Disease
; FILE REFERENCE: 0609 4460005
; CURRENT APPLICATION NUMBER: US/10/052,817
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: 09/241,606
; PRIOR FILING DATE: 1999-02-02
; PRIOR APPLICATION NUMBER: 09/148,503
; PRIOR FILING DATE: 1998-09-04
; PRIOR APPLICATION NUMBER: 60/093,297
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 6
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-052-817-6

```

```

Query Match          16.7%; Score 5; DB 12; Length 111;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      3 SLPBL 7
        |||||
Db      80 SLPBL 84

```

RESULT 29

```

US-09-846-590B-2
; Sequence 2, Application US/09846590B
; Patent No. US2002009801A1
; GENERAL INFORMATION:
; APPLICANT: Falco, S. Carl
; APPLICANT: Farnodu, Layo O.
; APPLICANT: Orozco, Emil M. Jr.
; TITLE OF INVENTION: Amino Acid Decarboxylases
; FILE REFERENCE: BB-1237
; CURRENT APPLICATION NUMBER: US/09/846,590B
; CURRENT FILING DATE: 2001-05-01
; PRIOR APPLICATION NUMBER: 60/099,493
; PRIOR FILING DATE: 1998-09-08
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 127
; TYPE: PRT
; ORGANISM: Catalpa speciosa
US-09-846-590B-2

```

```

Query Match          16.7%; Score 5; DB 10; Length 127;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      4 LPELD 8
        |||||
Db      48 LPELD 52

```

```

RESULT 30
US-09-164-615-13
; Sequence 13, Application US/09164615B
; Patent No. US20020073442A1
; GENERAL INFORMATION:
; APPLICANT: Hanley-Bowdoin, Linda
; APPLICANT: Settlage, Sharon
; TITLE OF INVENTION: Geminivirus Resistant Transgenic Plants
; FILE REFERENCE: 5051-433
; CURRENT APPLICATION NUMBER: US/09/164,615B
; CURRENT FILING DATE: 1998-10-01
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 13
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Pepper huasteco virus
US-09-164-615-13

```

```

Query Match          16.7%; Score 5; DB 10; Length 132;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      18 ISGOI 22
        |||||
Db      77 ISGOI 81

```

```

RESULT 31
US-10-101-464A-485
; Sequence 485, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01

```

```
/ PRIOR APPLICATION NUMBER: 09/228,986
/ PRIOR FILING DATE: 1999-01-12
/ PRIOR APPLICATION NUMBER: 60/162,866
/ PRIOR FILING DATE: 1999-11-01
/ PRIOR APPLICATION NUMBER: PCT/US00/00724
/ PRIOR FILING DATE: 2000-01-11
/ NUMBER OF SEQ ID NOS: 989
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 485
/ LENGTH: 133
/ TYPE: PRT
/ ORGANISM: Eucalyptus grandis
US-10-101-464A-485
```

```
Query Match      16.7% Score 5; DB 9; Length 133;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      4 LPELD 8
      |||||
      83 LPELD 87
```

```
RESULT 32
US-10-147-454-1
/ Sequence 1, Application US/10147454
/ Publication No. US20020193308A1
/ GENERAL INFORMATION:
/ APPLICANT: NAKTINIS, Vytautas
/ APPLICANT: CONCAS BENEVELLI, Daniela
/ APPLICANT: BERRA, Bruno
/ APPLICANT: COLOMBO, Irma
/ APPLICANT: RONCHI, Severino
/ APPLICANT: BARTORELLI, Alberto
/ APPLICANT: BUMELIS, Viadas Algirdus
/ TITLE OF INVENTION: Recombinant protein and its use in Therapy and Diagnostics
/ FILE REFERENCE: 0471-0273P
/ CURRENT APPLICATION NUMBER: US/10/147,454
/ CURRENT FILING DATE: 2002-05-17
/ NUMBER OF SEQ ID NOS: 9
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 1
/ LENGTH: 136
/ TYPE: PRT
/ ORGANISM: Capra hircus
US-10-147-454-1
```

```
Query Match      16.7% Score 5; DB 9; Length 136;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      17 YISGQ 21
      |||||
      DB      31 YISGQ 35
```

```
RESULT 33
US-10-147-454-3
/ Sequence 3, Application US/10147454
/ Publication No. US20020193308A1
/ GENERAL INFORMATION:
/ APPLICANT: NAKTINIS, Vytautas
/ APPLICANT: CONCAS BENEVELLI, Daniela
/ APPLICANT: BERRA, Bruno
/ APPLICANT: COLOMBO, Irma
/ APPLICANT: RONCHI, Severino
/ APPLICANT: BARTORELLI, Alberto
/ APPLICANT: BUMELIS, Viadas Algirdus
/ TITLE OF INVENTION: Recombinant Protein and its Use in Therapy and Diagnostics
/ FILE REFERENCE: 0471-0273P
/ CURRENT APPLICATION NUMBER: US/10/147,454
/ CURRENT FILING DATE: 2002-05-17
/ NUMBER OF SEQ ID NOS: 9
```

```
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 3
/ LENGTH: 137
/ TYPE: PRT
/ ORGANISM: Capra hircus
US-10-147-454-3
```

```
Query Match      16.7% Score 5; DB 9; Length 137;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      17 YISGQ 21
      |||||
      DB      32 YISGQ 36
```

```
RESULT 34
US-09-291-332-1
/ Sequence 1, Application US/09291332A
/ Patent No. US20010014471A1
/ GENERAL INFORMATION:
/ APPLICANT: NAKTINIS, Vytautas
/ APPLICANT: CONCAS BENEVELLI, Daniela
/ APPLICANT: BERRA, Bruno
/ APPLICANT: COLOMBO, Irma
/ APPLICANT: RONCHI, Severino
/ APPLICANT: BARTORELLI, Alberto
/ APPLICANT: BUMELIS, Viadas
/ TITLE OF INVENTION: Recombinant protein and its use in therapy and
/ FILE REFERENCE: zetes15
/ CURRENT APPLICATION NUMBER: US/09/291,332A
/ CURRENT FILING DATE: 1999-04-15
/ NUMBER OF SEQ ID NOS: 2
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 1
/ LENGTH: 137
/ TYPE: PRT
/ ORGANISM: Capra hircus
US-09-291-332-1
```

```
Query Match      16.7% Score 5; DB 10; Length 137;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      17 YISGQ 21
      |||||
      DB      32 YISGQ 36
```

```
RESULT 35
US-10-147-454-2
/ Sequence 2, Application US/10147454
/ Publication No. US20020193308A1
/ GENERAL INFORMATION:
/ APPLICANT: NAKTINIS, Vytautas
/ APPLICANT: CONCAS BENEVELLI, Daniela
/ APPLICANT: BERRA, Bruno
/ APPLICANT: COLOMBO, Irma
/ APPLICANT: RONCHI, Severino
/ APPLICANT: BARTORELLI, Alberto
/ APPLICANT: BUMELIS, Viadas Algirdus
/ TITLE OF INVENTION: Recombinant Protein and its Use in Therapy and Diagnostics
/ FILE REFERENCE: 0471-0273P
/ CURRENT APPLICATION NUMBER: US/10/147,454
/ CURRENT FILING DATE: 2002-05-17
/ NUMBER OF SEQ ID NOS: 9
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 2
/ LENGTH: 139
/ TYPE: PRT
/ ORGANISM: Capra hircus
US-10-147-454-2
```



Query Match 16.7%; Score 5; DB 9; Length 139;  
Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 17 YISGO 21  
DB 34 YISGO 38

## RESULT 36

US-09-738-626-3519  
; Sequence 3519, Application US/09738626  
; Publication No. US20020197605A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAGAWA, SATOSHI  
; APPLICANT: MIZOGUCHI, HIROSHI  
; APPLICANT: ANDO, SEIKO  
; APPLICANT: HAYASHI, MIKIRO  
; APPLICANT: OCHIAI, KEIKO  
; APPLICANT: YOKOI, HARUHIKO  
; APPLICANT: TATEISHI, NAOKO  
; APPLICANT: SENOH, AKIHIRO  
; APPLICANT: IKEDA, MASATO  
; APPLICANT: OZAKI, AKIO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-125  
; CURRENT APPLICATION NUMBER: US/09/738, 626  
; PRIOR FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: JP 99/377484  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: JP 00/159162  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: JP 00/280988  
; PRIOR FILING DATE: 2000-08-03  
; NUMBER OF SEQ ID NOS: 7059  
; SOFTWARE: PatentIn ver. 3.0  
; SEQ ID NO 3519  
; LENGTH: 140  
; TYPE: PRT  
; ORGANISM: Corynebacterium glutamicum  
US-09-738-626-3519

Query Match 16.7%; Score 5; DB 9; Length 140;  
Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

8 DYERS 12  
73 DYERS 77

## RESULT 37

US-10-094-214-4  
; Sequence 4, Application US/10094214  
; Patent No. US20020132303A1  
; GENERAL INFORMATION:  
; APPLICANT: Curtis, Rory A.J.  
; APPLICANT: Millennium Pharmaceuticals Inc.  
; TITLE OF INVENTION: (Transporter) Family Member and Uses Therefor  
; FILE REFERENCE: MP101-038P1RM  
; CURRENT APPLICATION NUMBER: US/10/094, 214  
; PRIOR FILING DATE: 2002-03-08  
; PRIOR APPLICATION NUMBER: 60/275, 078  
; PRIOR FILING DATE: 2001-03-12  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 152  
; TYPE: PRT  
; ORGANISM: Artificial Sequence

OTHER INFORMATION: consensus  
US-10-094-214-4

Query Match 16.7%; Score 5; DB 12; Length 152;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SLPEL 7  
DB 38 SLPEL 42

## RESULT 38

US-09-791-932-73  
; Sequence 73, Application US/09791932  
; Publication No. US20030003451A1  
; GENERAL INFORMATION:  
; APPLICANT: Vogeli, Gabriel  
; APPLICANT: Parodi, Luis A.  
; APPLICANT: Hiebsch, Ronald R.  
; APPLICANT: Lind, Peter  
; APPLICANT: Kaytes, Paul S.  
; APPLICANT: Ruff, Valerie  
; APPLICANT: Huff, Rita M.  
; APPLICANT: Wood, Linda S.  
; TITLE OF INVENTION: No. US20030003451A1 G Protein-Coupled Receptors Cross-Refere  
; FILE REFERENCE: 00325\_US1  
; CURRENT APPLICATION NUMBER: US/09/791, 932  
; PRIOR FILING DATE: 2001-02-23  
; PRIOR APPLICATION NUMBER: 60/184, 305  
; PRIOR FILING DATE: 2000-02-23  
; PRIOR APPLICATION NUMBER: 60/184, 304  
; PRIOR FILING DATE: 2000-02-23  
; PRIOR APPLICATION NUMBER: 60/184, 303  
; PRIOR FILING DATE: 2000-02-23  
; PRIOR APPLICATION NUMBER: 60/184, 397  
; PRIOR FILING DATE: 2000-02-23  
; PRIOR APPLICATION NUMBER: 60/184, 247  
; PRIOR FILING DATE: 2000-02-23  
; PRIOR APPLICATION NUMBER: 60/188, 880  
; PRIOR FILING DATE: 2000-03-13  
; PRIOR APPLICATION NUMBER: 60/217, 369  
; PRIOR FILING DATE: 2000-07-11  
; PRIOR APPLICATION NUMBER: 60/217, 370  
; PRIOR FILING DATE: 2000-07-11  
; PRIOR APPLICATION NUMBER: 60/218, 492  
; PRIOR FILING DATE: 2000-07-20  
; PRIOR APPLICATION NUMBER: 60/186, 810  
; PRIOR FILING DATE: 2000-03-03  
; PRIOR APPLICATION NUMBER: 60/188, 064  
; PRIOR FILING DATE: 2000-03-09  
; PRIOR APPLICATION NUMBER: 60/186, 457  
; PRIOR FILING DATE: 2000-03-02  
; PRIOR APPLICATION NUMBER: 60/213, 861  
; PRIOR FILING DATE: 2000-06-23  
; PRIOR APPLICATION NUMBER: 60/194, 344  
; PRIOR FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: 60/218, 337  
; PRIOR FILING DATE: 2000-07-14  
; NUMBER OF SEQ ID NOS: 184  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 73  
; LENGTH: 157  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-791-932-73

Query Match 16.7%; Score 5; DB 9; Length 157;  
Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SLPEL 7  
DB 38 SLPEL 42

Db 86 SLPBL 90

## RESULT 39

US-10-112-793-16

; Sequence 16, Application US/10112793

; Publication No. US20020192729A1

; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi J.

; TITLE OF INVENTION: Apo-2 LI AND Apo-3 POLYPEPTIDES

; NUMBER OF SEQUENCES: 28

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Genentech, Inc.

; STREET: 1 DNA Way

; CITY: South San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94080

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: WinPattn (Genentech)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/112,793

; FILING DATE: 28-Mar-2002

; CLASSIFICATION: &lt;Unknown&gt;

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/828,683A

; FILING DATE: 31-Mar-1997

; APPLICATION NUMBER: 08/625328

; FILING DATE: 1-Apr-1996

; APPLICATION NUMBER: 08/710802

; FILING DATE: 23-Sep-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Marschang, Diane L.

; REGISTRATION NUMBER: 35,600

; REFERENCE/DOCKET NUMBER: P1007P1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 650/225-5416

; TELEFAX: 650/952-9881

; INFORMATION FOR SEQ ID NO: 16:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 159 amino acids

; TYPE: Amino Acid

; TOPOLOGY: Linear

; SEQUENCE DESCRIPTION: SEQ ID NO: 16:

US-10-112-793-16

## Query Match

Best Local Similarity 100.0%; Score 5; DB 9; Length 159;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 SATEP 16

Db 48 SATEP 52

## RESULT 40

US-10-101-464A-650

; Sequence 650, Application US/10101464A

; Publication No. US20030046728A1

; GENERAL INFORMATION:

; APPLICANT: Strabala, Timothy

; APPLICANT: Nieuwenhuizen, Nicolaas

; APPLICANT: Higgins, Colleen M.

; TITLE OF INVENTION: Compositions Isolated from Plant Cells

; FILE REFERENCE: 11000.1020C2

; CURRENT APPLICATION NUMBER: US/10/101,464A

; FILING DATE: 2002-03-18

; PRIOR APPLICATION NUMBER: 09/704,302

; PRIOR FILING DATE: 2000-11-01

; PRIOR APPLICATION NUMBER: 09/228,986

; PRIOR FILING DATE: 1999-01-12

; PRIOR APPLICATION NUMBER: 60/162,866

; PRIOR FILING DATE: 1999-11-01

; PRIOR APPLICATION NUMBER: PCT/US00/00724

; PRIOR FILING DATE: 2000-01-11

; NUMBER OF SEQ ID NOS: 989

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 650

; LENGTH: 159

; TYPE: PRT

; ORGANISM: Eucalyptus grandis

US-10-101-464A-650

## Query Match

Best Local Similarity 100.0%; Score 5; DB 9; Length 159;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SLPBL 7

Db 87 SLPBL 91

Search completed: April 9, 2003, 14:29:02  
Job time : 27.055 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 9, 2003, 13:59:59 ; Search time 6.33028 Seconds  
(without alignments)  
289.732 Million cell updates/sec

Title: US-09-987-190-2

Perfect score: 145  
Sequence: 1 KYSLPELDYFSATEPYISQINEIXYTX 30

Scoring table: BLOSUM62  
Gap 10.0 , Gapext 0.5

Searched: 248812 seqs, 61136040 residues

Actual number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications\_AA:\*  
1: /cgn2\_6/ptodata/2/pubppa/US08\_NEW\_PUB pep:\*  
2: /cgn2\_6/ptodata/2/pubppa/PCT\_NEW\_PUB pep:\*  
3: /cgn2\_6/ptodata/2/pubppa/US06\_NEW\_PUB pep:\*  
4: /cgn2\_6/ptodata/2/pubppa/US06\_PUBCOMB pep:\*  
5: /cgn2\_6/ptodata/2/pubppa/US07\_NEW\_PUB pep:\*  
6: /cgn2\_6/ptodata/2/pubppa/US07\_PUBCOMB pep:\*  
7: /cgn2\_6/ptodata/2/pubppa/PCTUS\_PUBCOMB pep:\*  
8: /cgn2\_6/ptodata/2/pubppa/US08\_PUBCOMB pep:\*  
9: /cgn2\_6/ptodata/2/pubppa/US09\_NEW\_PUB pep:\*  
10: /cgn2\_6/ptodata/2/pubppa/US09\_PUBCOMB pep:\*  
11: /cgn2\_6/ptodata/2/pubppa/US10\_NEW\_PUB pep:\*  
12: /cgn2\_6/ptodata/2/pubppa/US10\_PUBCOMB pep:\*  
13: /cgn2\_6/ptodata/2/pubppa/US60\_NEW\_PUB pep:\*  
14: /cgn2\_6/ptodata/2/pubppa/US60\_PUBCOMB pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	141	97.2	30	10 US-09-987-190-2	Sequence 2, Appl
2	77	53.1	200	9 US-09-738-626-6727	Sequence 6727, Ap
3	77	53.1	200	10 US-09-818-564-2	Sequence 12, Appl
4	74	51.0	95	9 US-09-993-333-12	Sequence 6, Appl
5	71	49.0	188	10 US-09-987-190-6	Sequence 7, Appl
6	67	46.2	198	10 US-09-727-8558-7	Sequence 23, Appl
7	66	45.5	22	9 US-09-953-510-23	Sequence 974, App
8	66	45.5	222	9 US-09-727-8558-5	Sequence 974, App
9	50	34.5	359	9 US-09-895-793-974	Sequence 708, App
10	50	34.5	371	10 US-09-822-827-974	Sequence 708, App
11	50	34.5	371	9 US-10-012-896-708	Sequence 708, App
12	50	34.5	371	9 US-09-895-793-708	Sequence 708, App
13	50	34.5	371	9 US-09-895-814-708	Sequence 708, App
14	50	34.5	371	10 US-09-759-143-708	Sequence 708, App
15	50	34.5	371	10 US-09-780-669-708	Sequence 708, App
16	50	34.5	371	10 US-09-822-827-708	Sequence 708, App
17	50	34.5	400	9 US-10-012-896-852	Sequence 852, App
18	50	34.5	400	9 US-09-895-793-852	Sequence 852, App
19	50	34.5	400	9 US-09-895-814-852	Sequence 852, App

20	50	34.5	400	10 US-09-759-143-852	Sequence 852, App
21	50	34.5	400	10 US-09-780-669-852	Sequence 852, App
22	50	34.5	400	10 US-09-822-827-852	Sequence 113, App
23	50	34.5	553	9 US-09-232-880-113	Sequence 113, App
24	50	34.5	553	9 US-10-012-896-113	Sequence 113, App
25	50	34.5	553	9 US-09-895-793-113	Sequence 113, App
26	50	34.5	553	9 US-09-895-814-113	Sequence 113, App
27	50	34.5	553	10 US-09-745-288-101	Sequence 101, App
28	50	34.5	553	10 US-09-838-785-2	Sequence 113, App
29	50	34.5	553	10 US-09-759-143-113	Sequence 113, App
30	50	34.5	553	10 US-09-780-669-113	Sequence 113, App
31	50	34.5	553	10 US-09-030-606-113	Sequence 113, App
32	50	34.5	553	10 US-09-822-827-113	Sequence 113, App
33	50	34.5	553	10 US-09-115-453-113	Sequence 113, App
34	50	34.5	1079	9 US-09-895-793-947	Sequence 947, App
35	50	34.5	1079	10 US-09-822-827-947	Sequence 947, App
36	45.5	31.4	203	10 US-09-972-484-11	Sequence 11, Appl
37	45.5	31.4	1101	10 US-09-972-484-52	Sequence 52, Appl
38	45.5	31.0	1452	10 US-09-972-484-2	Sequence 122, App
39	45	31.0	280	9 US-09-832-355A-122	Sequence 96, Appl
40	45	31.0	360	9 US-09-832-355A-96	Sequence 5, Appl
41	45	31.0	495	10 US-09-998-831-5	Sequence 6, Appl
42	45	31.0	496	9 US-10-179-744-6	Sequence 6, Appl
43	45	31.0	496	9 US-10-186-817-6	Sequence 6, Appl
44	45	31.0	496	9 US-10-215-224-8	Sequence 8, Appl
45	45	31.0	496	10 US-09-998-831-4	Sequence 4, Appl

#### ALIGNMENTS

```
RESULT 1
US-09-987-190-2
; Sequence 2, Application US/09987190
; Patent No. US20020058293A1
; GENERAL INFORMATION:
; APPLICANT: TAKESAKO, Kazutoh
; APPLICANT: MIZUTANI, Shigetoshi
; APPLICANT: ENDO, Masahiro
; APPLICANT: KATO, Kunoshin
; TITLE OF INVENTION: FUNGAL ANTIGENS AND PROCESS FOR PRODUCING THE SAME
; FILE REFERENCE: 1422-0502P
; CURRENT APPLICATION NUMBER: US/09/987,190
; PRIOR APPLICATION NUMBER: 09/262,856
; PRIOR FILING DATE: 1999-03-04
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: Unsure
; LOCATION: (1)..(30)
; OTHER INFORMATION: any Xaa = any amino acid, unknown or other
US-09-987-190-2

Query Match      97.2%  Score 141,  DB 10,  Length 30;
Best Local Similarity 100.0%;  Pred. No. 3.9e-15;
Matches 28;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Cy      1  KYSLPELDYFSATEPYISQINEIXYT 28
Db      1  KYSLPELDYFSATEPYISQINEIXYT 28

RESULT 2
US-09-738-626-6727
; Sequence 6727, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
```

APPLICANT: MIZOGUCHI, HIROSHI  
APPLICANT: ANDO, SEIKO  
APPLICANT: HAYASHI, MIKIRO  
APPLICANT: OCHIAI, KEIKO  
APPLICANT: YOKOI, HARUHIKO  
APPLICANT: TATEISHI, NAOKO  
APPLICANT: SENOH, AKIHIRO  
APPLICANT: IKEDA, MASATO  
APPLICANT: OZAKI, AKIO  
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
FILE REFERENCE: 249-125  
CURRENT APPLICATION NUMBER: US/09/738,626  
CURRENT FILING DATE: 2000-12-18  
PRIOR APPLICATION NUMBER: JP 99/377484  
PRIOR FILING DATE: 1999-12-16  
PRIOR APPLICATION NUMBER: JP 00/159162  
PRIOR FILING DATE: 2000-04-07  
PRIOR APPLICATION NUMBER: JP 00/280988  
PRIOR FILING DATE: 2000-08-03  
NUMBER OF SEQ ID NOS: 7059  
SOFTWARE: PatentIn ver. 3.0  
SEQ ID NO 6727  
LENGTH: 200  
TYPE: PRT  
ORGANISM: Corynebacterium glutamicum  
US-09-738-626-6727

Query Match 53.1%; Score 77; DB 9; Length 200;  
Best Local Similarity 48.1%; Pred. No. 0.00022;  
Matches 13; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 2 YSLPELDYFSAATEPYISQINEIXYT 28  
DB 4 YELPELDYADALBPHIAEIMELHHS 30

RESULT 3  
US-09-818-564-2  
Sequence 2, Application US/09818564  
Patent No. US20020137151A1  
GENERAL INFORMATION:  
APPLICANT: MUKOMAM, MURIEL  
APPLICANT: GUYONVARCH, ARMEL  
APPLICANT: MAXX, ACHIM  
TITLE OF INVENTION: A PROCESS FOR THE FERMENTATIVE PREPARATION OF METABOLIC  
TITLE OF INVENTION: PRODUCTS AND FOR THE NUCLEOTIDE SEQUENCES ENCODING FOR  
TITLE OF INVENTION: THE GOD GENE  
FILE REFERENCE: 21123/278411/MAS  
CURRENT APPLICATION NUMBER: US/09/818,564  
CURRENT FILING DATE: 2001-03-28  
PRIOR APPLICATION NUMBER: 09/373,731  
PRIOR FILING DATE: 1999-08-13  
NUMBER OF SEQ ID NOS: 8  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 200  
TYPE: PRT  
ORGANISM: Corynebacterium melassaccola  
OTHER INFORMATION: ATCC 17965  
US-09-818-564-2

Query Match 53.1%; Score 77; DB 10; Length 200;  
Best Local Similarity 48.1%; Pred. No. 0.00022;  
Matches 13; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 2 YSLPELDYFSAATEPYISQINEIXYT 28  
DB 4 YELPELDYADALBPHIAEIMELHHS 30

RESULT 4  
US-09-993-333-12  
Sequence 12, Application US/09993333

Patent No. US20020156040A1  
GENERAL INFORMATION:  
APPLICANT: Oberley, Larry Wayne  
APPLICANT: Meydert, Christine J.  
APPLICANT: Smith, Benjamin Barnes  
TITLE OF INVENTION: Reduction of antioxidant enzyme levels in tumor cells using anti  
TITLE OF INVENTION: Oligonucleotides  
FILE REFERENCE: 875,042US1  
CURRENT APPLICATION NUMBER: US/09/993,333  
CURRENT FILING DATE: 2001-11-14  
PRIOR APPLICATION NUMBER: US 60/248,328  
PRIOR FILING DATE: 2000-11-14  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 12  
LENGTH: 95  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-993-333-12

Query Match 51.0%; Score 74; DB 9; Length 95;  
Best Local Similarity 42.9%; Pred. No. 0.00027;  
Matches 12; Conservative 10; Mismatches 6; Indels 0; Gaps 0;

QY 1 KXSLPELDYFSAATEPYISQINEIXYT 28  
DB 25 KXSLPELDYADALBPHIAEIMELHHS 52

RESULT 5  
US-09-987-190-6  
Sequence 6, Application US/09987190  
Patent No. US20020058233A1  
GENERAL INFORMATION:  
APPLICANT: TAKESAKO, Kazutosh  
APPLICANT: MIZUTANI, Shigetoshi  
APPLICANT: ENDO, Masahiro  
APPLICANT: KATO, Ikunoshin  
TITLE OF INVENTION: FUNGAL ANTIGENS AND PROCESS FOR PRODUCING THE SAME  
FILE REFERENCE: 1422-0502P  
CURRENT APPLICATION NUMBER: US/09/987,190  
CURRENT FILING DATE: 2001-11-13  
PRIOR APPLICATION NUMBER: 09/262,856  
PRIOR FILING DATE: 1999-03-04  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 6  
LENGTH: 188  
TYPE: PRT  
ORGANISM: Candida albicans  
US-09-987-190-6

Query Match 49.0%; Score 71; DB 10; Length 188;  
Best Local Similarity 81.2%; Pred. No. 0.0017;  
Matches 13; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 13 ATEPYISQINEIXYT 28  
DB 1 ATEPYISQINEIXYT 16

RESULT 6  
US-09-727-855B-7  
Sequence 7, Application US/09727855B  
Patent No. US20020168703A1  
GENERAL INFORMATION:  
APPLICANT: HOSHINO, Tatsuo  
APPLICANT: OJIMA, Kazuyuki  
APPLICANT: SETOGUCHI, Yutaka  
TITLE OF INVENTION: PROCESS FOR THE MANUFACTURE OF CAROTENOIDS AND BIOLOGICALLY USEF  
FILE REFERENCE: C38455/111694  
CURRENT APPLICATION NUMBER: US/09/727,855B



```
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 974
/ LENGTH: 359
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-895-793-974
```

```
Query Match          34.5%; Score 50; DB 9; Length 359;
Best Local Similarity 45.0%; Pred. No. 5.7;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
```

```
QY      2 YSLPDLDFEFSATEPYISQ 21
Db      176 YLLPALDMDTSLALAPYLCTQ 195
```

```
RESULT 10
US-09-822-827-974
/ Sequence 974, Application US/09822827
/ Patent No. US20020081680A1
GENERAL INFORMATION:
```

```
APPLICANT: Xu, Jiangchun
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ FILE REFERENCE: 210121.534C1
/ CURRENT APPLICATION NUMBER: US/09/822,827
/ CURRENT FILING DATE: 2001-03-28
/ NUMBER OF SEQ ID NOS: 982
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 974
/ LENGTH: 359
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-822-827-974
```

```
Query Match          34.5%; Score 50; DB 10; Length 359;
Best Local Similarity 45.0%; Pred. No. 5.7;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
```

```
QY      2 YSLPDLDFEFSATEPYISQ 21
Db      176 YLLPALDMDTSLALAPYLCTQ 195
```

```
RESULT 11
US-10-012-896-708
```

```
/ Sequence 708, Application US/10012896
/ Publication No. US20020183251A1
GENERAL INFORMATION:
```

```
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yugu
APPLICANT: Kalos, Michael D.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William T.
APPLICANT: Henderson, Robert A.
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghton, Raymond L.
APPLICANT: Vinals de Bassols, Carlota
APPLICANT: Foy, Teresa
APPLICANT: Fanger, Gary R.
APPLICANT: Mantanabe, Yoshihiro
APPLICANT: Meagher, Madeleine Joy
```

```
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ FILE REFERENCE: 210121.427C27
/ CURRENT APPLICATION NUMBER: US/10/012,896
/ CURRENT FILING DATE: 2001-12-10
/ NUMBER OF SEQ ID NOS: 1011
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 708
/ LENGTH: 371
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-012-896-708
```

```
Query Match          34.5%; Score 50; DB 9; Length 371;
Best Local Similarity 45.0%; Pred. No. 5.9;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
```

```
QY      2 YSLPDLDFEFSATEPYISQ 21
Db      150 YLLPALDMDTSLALAPYLCTQ 169
```

```
RESULT 12
US-09-895-793-708
/ Sequence 708, Application US/09895793
/ Publication No. US20020192763A1
GENERAL INFORMATION:
```

```
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yugu
APPLICANT: Kalos, Michael D.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William T.
APPLICANT: Henderson, Robert A.
APPLICANT: Hural, John
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghton, Raymond L.
APPLICANT: Vinals de Bassols, Carlota
APPLICANT: Foy, Teresa
APPLICANT: Fanger, Gary R.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ FILE REFERENCE: 210121.534C2
/ CURRENT APPLICATION NUMBER: US/09/895,793
/ CURRENT FILING DATE: 2001-06-29
/ NUMBER OF SEQ ID NOS: 982
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 708
/ LENGTH: 371
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-895-793-708
```

```
Query Match          34.5%; Score 50; DB 9; Length 371;
Best Local Similarity 45.0%; Pred. No. 5.9;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
```

```
QY      2 YSLPDLDFEFSATEPYISQ 21
Db      150 YLLPALDMDTSLALAPYLCTQ 169
```

```
RESULT 13
US-09-895-814-708
```

```
; Sequence 708, Application US/09895814
; Publication No. US20020193296A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlotca
; APPLICANT: Roy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C26
; CURRENT APPLICATION NUMBER: US/09/895.814
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 708
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-895-814-708

Query Match          34.5%; Score 50; DB 9; Length 371;
Best Local Similarity 45.0%; Pred. No. 5.9;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY      2 YSLPELDYFSGATEPYISGQ 21
      |||::|||::|
Db      150 YLPAIDWDTSALAPYIGTQ 169

SULT 14
-09-759-143-708
; Sequence 708, Application US/09759143
; Patent No. US20020022248A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C23
; CURRENT APPLICATION NUMBER: US/09/759.143
```

```
; CURRENT FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 708
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-759-143-708

Query Match          34.5%; Score 50; DB 10; Length 371;
Best Local Similarity 45.0%; Pred. No. 5.9;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY      2 YSLPELDYFSGATEPYISGQ 21
      |||::|||::|
Db      150 YLPAIDWDTSALAPYIGTQ 169

RESULT 15
US-09-780-669-708
; Sequence 708, Application US/09780669
; Patent No. US20020051977A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C24
; CURRENT APPLICATION NUMBER: US/09/780.669
; CURRENT FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 943
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 708
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-780-669-708

Query Match          34.5%; Score 50; DB 10; Length 371;
Best Local Similarity 45.0%; Pred. No. 5.9;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY      2 YSLPELDYFSGATEPYISGQ 21
      |||::|||::|
Db      150 YLPAIDWDTSALAPYIGTQ 169

RESULT 16
US-09-822-827-708
; Sequence 708, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
```

```
/ FILE REFERENCE: 210121.534C1
/ CURRENT APPLICATION NUMBER: US/09/822.827
/ CURRENT FILING DATE: 2001-03-28
/ NUMBER OF SEQ ID NOS: 982
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 708
/ LENGTH: 371
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-822-827-708
```

```
Query Match      34.5%; Score 50; DB 10; Length 371;
Best Local Similarity 45.0%; Pred. No. 5.9;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
```

```
QY      2 YSLPDLDFEFSATPEYISGQ 21
      |||::|||::|||
Db      150 YLLPAIDWDTSLAPYLGCTQ 169
```

```
SULT 17
-10-012-896-852
/ Sequence 852, Application US/100128936
/ Publication No. US20020183251A1
/ GENERAL INFORMATION:
/ APPLICANT: Xu, Jiangchun
/ APPLICANT: Dillon, Davin C.
/ APPLICANT: Mitcham, Jennifer L.
/ APPLICANT: Harlocker, Susan L.
/ APPLICANT: Jiang, Yugu
/ APPLICANT: Kalos, Michael D.
/ APPLICANT: Retter, Marc W.
/ APPLICANT: Stolk, John A.
/ APPLICANT: Day, Craig H.
/ APPLICANT: Vedvick, Thomas S.
/ APPLICANT: Carter, Darrick
/ APPLICANT: Li, Samuel X.
/ APPLICANT: Wang, Aijun
/ APPLICANT: Skeiky, Yasir A.W.
/ APPLICANT: Heppler, William T.
/ APPLICANT: Henderson, Robert A.
/ APPLICANT: Hurai, John
/ APPLICANT: McNeill, Patricia D.
/ APPLICANT: Houghton, Raymond L.
/ APPLICANT: Vinals de Bassols, Carlota
/ APPLICANT: Foy, Teresa
/ APPLICANT: Fanger, Gary R.
/ APPLICANT: Wantanabe, Yoshihito
/ APPLICANT: Meesher, Madeleine Joy
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ FILE REFERENCE: 210121.427C27
/ CURRENT APPLICATION NUMBER: US/10/012.896
/ CURRENT FILING DATE: 2001-12-10
/ NUMBER OF SEQ ID NOS: 1011
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 852
/ LENGTH: 400
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-012-896-852
```

```
Query Match      34.5%; Score 50; DB 9; Length 400;
Best Local Similarity 45.0%; Pred. No. 6.5;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
```

```
QY      2 YSLPDLDFEFSATPEYISGQ 21
      |||::|||::|||
Db      278 YLLPAIDWDTSLAPYLGCTQ 297
```

```
RESULT 18
US-09-895-793-852
```

```
/ Sequence 852, Application US/09895793
/ Publication No. US20020192763A1
/ GENERAL INFORMATION:
/ APPLICANT: Xu, Jiangchun
/ APPLICANT: Dillon, Davin C.
/ APPLICANT: Mitcham, Jennifer L.
/ APPLICANT: Harlocker, Susan L.
/ APPLICANT: Jiang, Yugu
/ APPLICANT: Kalos, Michael D.
/ APPLICANT: Retter, Marc W.
/ APPLICANT: Stolk, John A.
/ APPLICANT: Day, Craig H.
/ APPLICANT: Vedvick, Thomas S.
/ APPLICANT: Carter, Darrick
/ APPLICANT: Li, Samuel X.
/ APPLICANT: Wang, Aijun
/ APPLICANT: Skeiky, Yasir A.W.
/ APPLICANT: Heppler, William T.
/ APPLICANT: Henderson, Robert A.
/ APPLICANT: Hurai, John
/ APPLICANT: McNeill, Patricia D.
/ APPLICANT: Houghton, Raymond L.
/ APPLICANT: Vinals de Bassols, Carlota
/ APPLICANT: Foy, Teresa
/ APPLICANT: Fanger, Gary R.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ FILE REFERENCE: 210121.534C2
/ CURRENT APPLICATION NUMBER: US/09/895.793
/ CURRENT FILING DATE: 2001-06-29
/ NUMBER OF SEQ ID NOS: 982
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 852
/ LENGTH: 400
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-895-793-852
```

```
Query Match      34.5%; Score 50; DB 9; Length 400;
Best Local Similarity 45.0%; Pred. No. 6.5;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
```

```
QY      2 YSLPDLDFEFSATPEYISGQ 21
      |||::|||::|||
Db      278 YLLPAIDWDTSLAPYLGCTQ 297
```

```
RESULT 19
US-09-895-814-852
/ Sequence 852, Application US/09895814
/ Publication No. US20020193296A1
/ GENERAL INFORMATION:
/ APPLICANT: Xu, Jiangchun
/ APPLICANT: Dillon, Davin C.
/ APPLICANT: Mitcham, Jennifer L.
/ APPLICANT: Harlocker, Susan L.
/ APPLICANT: Jiang, Yugu
/ APPLICANT: Kalos, Michael D.
/ APPLICANT: Retter, Marc W.
/ APPLICANT: Stolk, John A.
/ APPLICANT: Day, Craig H.
/ APPLICANT: Vedvick, Thomas S.
/ APPLICANT: Carter, Darrick
/ APPLICANT: Li, Samuel X.
/ APPLICANT: Wang, Aijun
/ APPLICANT: Skeiky, Yasir A.W.
/ APPLICANT: Heppler, William T.
/ APPLICANT: Henderson, Robert A.
/ APPLICANT: Hurai, John
/ APPLICANT: McNeill, Patricia D.
/ APPLICANT: Houghton, Raymond L.
/ APPLICANT: Vinals de Bassols, Carlota
/ APPLICANT: Foy, Teresa
```



```

: APPLICANT: Fanger, Gary R.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
: TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
: FILE REFERENCE: 210121.427C26
: CURRENT APPLICATION NUMBER: US/09/895,814
: CURRENT FILING DATE: 2001-06-29
: NUMBER OF SEQ ID NOS: 990
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 852
: LENGTH: 400
: TYPE: PR1
: ORGANISM: Homo sapiens
: US-09-895-814-852

```

Query Match	34.5%;	Score 50;	DB 9;	Length 400;
Best Local Similarity	45.0%;	Pred. No. 6.5;		
Matches	9;	Conservative	4;	Mismatches 7;
				Indels 0;
				Gaps 0;

```

2 YSLPELDYEFSAATEPYISGQ 21
  |||:|:|:|:|
278 YLLPAIDWDTSALAPYLGTQ 297

```

RESULT 20  
US-09-759-143-852  
; Sequence 852, Application US/09759143  
; Patent No. US20020022248A1

APPLICANT:	Xu, Jiangchun
APPLICANT:	Dillon, Devin C.
APPLICANT:	Matcham, Jennifer L.
APPLICANT:	Harlocker, Susan L.
APPLICANT:	Jiang, Yuqun
APPLICANT:	Henderson, Robert A.
APPLICANT:	Kalos, Michael D.
APPLICANT:	Fanger, Gary R.
APPLICANT:	Retter, Marc W.
APPLICANT:	Stolk, John A.
APPLICANT:	Day, Craig H.
APPLICANT:	Vedvick, Thomas S.
APPLICANT:	Carter, Darrick
APPLICANT:	Li, Samuel
APPLICANT:	Wang, Ajun
APPLICANT:	Skeiky, Yasir A. W.
APPLICANT:	Hepler, William
TITLE OF INVENTION:	DIAGNOSIS OF PROSTATE CANCER
TITLE OF INVENTION:	COMPOSITIONS AND METHODS FOR THE THERAPY AND

```
Query Match      34.5%; Score 50; DB 10; Length 400;
Best Local Similarity 45.0%; Pred. No. 6.5;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
```

```

Qy      2 YSLPELDYEFSAEPYISGQ 21
          |||:::||:|
Db      278 YLLPAIDWDTALAPYLGTQ 297

```

RESULT 21  
US-09-780-668-852  
; Sequence 852, Application US/09780665  
; Patent No. US20020051977A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun

```

APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Uiang, Yigui
APPLICANT: Henderson, Robert A.
APPLICANT: Kalos, Michael D.
APPLICANT: Fanger, Gary R.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedgwick, Thomas S.
APPLICANT: Carter, Derrick
APPLICANT: Li, Samuel
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yassar A.W.
APPLICANT: Hegler, William
APPLICANT: Hural, John
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghdon, Raymond L.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.427C24
CURRENT APPLICATION NUMBER: US/09/780,669
CURRENT FILING DATE: 2001-02-09
NUMBER OF SEQ ID NOS: 943
SOFTWARE: FastSeq Version 3.0
SEQ ID NO 852
LENGTH: 400
TYPE: PRT
ORGANISM: Homo sapiens
US-09-780-669--852

```

Query Match	34.5%	Score 50;	DB 10;	Length 400;
Best Local Similarity	45.0%	Pred. No. 6.5;		
Matches	9;	Conservative	4;	Mismatches 7; Indels 0; Gaps 0;

```
QY      2 YSLPELDYEFSAPEYISQ 21
        ||| :||| |||
Db      278 YLLPAIDWDTALAPYLGTQ 297
```

```

RESULT 22
US-09-822-827-852
: Sequence 852. Application US/09822827
: Patent No. US20020081680A1
: GENERAL INFORMATION:
: APPLICANT: Xu, Jiangchun
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
: TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
: FILE REFERENCE: 210121.534C1
: CURRENT APPLICATION NUMBER: US/09/822.827
: CURRENT FILING DATE: 2001-03-28
: NUMBER OF SEQ ID NOS: 982
: SOFTWARE: FASTSEQ for Windows Version 3.0
: SEQ ID NO 852
: LENGTH: 400
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-822-827-852

```

Query Match	34.5%	Score 50;	DB 10;	Length 400;
Best Local Similarity	45.0%;	Pred. No. 6.5;		
Matches	9;	Conservative	4;	Mismatches 7;
				Indels 0;
				Gaps 0;

```
QY      2 YSLPELDYEFSAEPYISGQ 21
      ||| : : : ||| :
Db      278 YLLPAIDWDTALAPYLGTQ 297
```

RESULT 23  
US-09-232-880-113  
; Sequence 113, Application US/09232880  
; Publication No. US20020182596A1

```
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer Lynn
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNODIAGNOSIS OF
; FILE REFERENCE: 210121.428C6
; CURRENT APPLICATION NUMBER: US/09/232,880
; NUMBER OF SEQ ID NOS: 338
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 113
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-232-880-113

Query Match      34.5%; Score 50; DB 9; Length 553;
Best Local Similarity 45.0%; Pred. No. 9.3;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

2 YSLPELDYFSGATEPYISQ 21
|||:|:|:|:|:|
Db 176 YLLPAIDWTSALAPYLGTD 195

RESULT 24
US-10-012-896-113
; Sequence 113, Application US/10012896
; Publication No. US20020183251A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darriek
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Heppler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; APPLICANT: Mantanabe, Yoshihiro
; APPLICANT: Meagher, Madeleine Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C27
; CURRENT APPLICATION NUMBER: US/10/012,896
; NUMBER OF SEQ ID NOS: 1011
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 113
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-012-896-113

Query Match      34.5%; Score 50; DB 9; Length 553;
Best Local Similarity 45.0%; Pred. No. 9.3;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

2 YSLPELDYFSGATEPYISQ 21
```

```
Db 176 YLLPAIDWTSALAPYLGTD 195

RESULT 25
US-09-895-793-113
; Sequence 113, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darriek
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Heppler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.534C2
; CURRENT APPLICATION NUMBER: US/09/895,793
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 113
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-895-793-113

Query Match      34.5%; Score 50; DB 9; Length 553;
Best Local Similarity 45.0%; Pred. No. 9.3;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

2 YSLPELDYFSGATEPYISQ 21
|||:|:|:|:|:|
Db 176 YLLPAIDWTSALAPYLGTD 195

RESULT 26
US-09-895-814-113
; Sequence 113, Application US/09895814
; Publication No. US20020193286A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darriek
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Heppler, William T.
```

```

; APPLICANT: Henderson, Robert A.
; APPLICANT: Hurai, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C26
; CURRENT APPLICATION NUMBER: US/09/895,814
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 113
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-895-814-113

```

```

Query Match          34.5%; Score 50; DB 9; Length 553;
Best Local Similarity 45.0%; Pred. No. 9.3;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

```

```

QY 2 YSLPELDYFSATEPYISQ 21
Db 176 YLPAIDWTSALAPYIGTQ 195

```

```

RESULT 27
US-09-745-288-101
; Sequence 101, Application US/09745288
; Patent No. US20010018058A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; FILE REFERENCE: 210121.446G1
; CURRENT APPLICATION NUMBER: US/09/745,288
; CURRENT FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 101
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapien
-09-745-288-101

```

```

Query Match          34.5%; Score 50; DB 10; Length 553;
Best Local Similarity 45.0%; Pred. No. 9.3;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

```

```

QY 2 YSLPELDYFSATEPYISQ 21
Db 176 YLPAIDWTSALAPYIGTQ 195

```

```

RESULT 28
US-09-838-785-2
; Sequence 2, Application US/09838785
; Patent No. US20020009455A1
; GENERAL INFORMATION:
; APPLICANT: Lau, Ted
; APPLICANT: Lin, Rick
; APPLICANT: Parkes, Debbie
; APPLICANT: Parry, Gordon
; APPLICANT: Schneider, Douglas
; APPLICANT: Steindricher, Renate
; APPLICANT: Van Heut, Pam T
; APPLICANT: Wu, John
; TITLE OF INVENTION: DNA Encoding a No. US20020009455A1e1 PROST 03

```

```

; FILE REFERENCE: 51831AUSM1
; CURRENT APPLICATION NUMBER: US/09/838,785
; CURRENT FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/200,065
; PRIOR FILING DATE: 2000-04-27
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-838-785-2

```

```

Query Match          34.5%; Score 50; DB 10; Length 553;
Best Local Similarity 45.0%; Pred. No. 9.3;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

```

```

QY 2 YSLPELDYFSATEPYISQ 21
Db 176 YLPAIDWTSALAPYIGTQ 195

```

```

RESULT 29
US-09-759-143-113
; Sequence 113, Application US/09759143
; Patent No. US2002002248A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqun
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C23
; CURRENT APPLICATION NUMBER: US/09/759,143
; CURRENT FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 113
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-759-143-113

```

```

Query Match          34.5%; Score 50; DB 10; Length 553;
Best Local Similarity 45.0%; Pred. No. 9.3;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

```

```

QY 2 YSLPELDYFSATEPYISQ 21
Db 176 YLPAIDWTSALAPYIGTQ 195

```

```

RESULT 30
US-09-780-669-113
; Sequence 113, Application US/09780669
; Patent No. US20020051977A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.

```

APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Jiang, Yugu  
APPLICANT: Henderson, Robert A.  
APPLICANT: Kalos, Michael D.  
APPLICANT: Fanger, Gary R.  
APPLICANT: Retter, Marc W.  
APPLICANT: Stoik, John A.  
APPLICANT: Day, Craig H.  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Carter, Derrick  
APPLICANT: Li, Samuel  
APPLICANT: Wang, Aijun  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Hepler, William  
APPLICANT: Hural, John  
APPLICANT: McNeill, Patricia D.  
APPLICANT: Houghton, Raymond L.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
FILE REFERENCE: 210121.427C24  
CURRENT APPLICATION NUMBER: US/09/780,669  
CURRENT FILING DATE: 2001-02-09  
NUMBER OF SEQ ID NOS: 943  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 113  
LENGTH: 553  
TYPE: PRT  
ORGANISM: Homo sapien  
US-09-780-669-113

Query Match 34.5%; Score 50; DB 10; Length 553;  
Best Local Similarity 45.0%; Pred. No. 9.3;  
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;  
QY 2 YSLPELDYFSATEPYISGQ 21  
DB 176 YLPAIDWTSALAPYLGTQ 195

RESULT 31  
US-09-030-606-113  
Sequence 113, Application US/09030606  
Patent No. US20020081580A1  
GENERAL INFORMATION:  
APPLICANT: Xu, Jiangchun  
APPLICANT: Dillon, Davin C.  
TITLE OF INVENTION: COMPOUNDS FOR IMMUNODIAGNOSIS OF PROSTATE CANCER AND METHODS F  
NUMBER OF SEQUENCES: 224  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/030,606  
FILING DATE: 25-FEB-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Makl, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.428C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 113:

SEQUENCE CHARACTERISTICS:  
LENGTH: 553 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
US-09-030-606-113

Query Match 34.5%; Score 50; DB 10; Length 553;  
Best Local Similarity 45.0%; Pred. No. 9.3;  
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;  
QY 2 YSLPELDYFSATEPYISGQ 21  
DB 176 YLPAIDWTSALAPYLGTQ 195

RESULT 32  
US-09-822-827-113  
Sequence 113, Application US/09822827  
Patent No. US20020081580A1  
GENERAL INFORMATION:  
APPLICANT: Xu, Jiangchun  
APPLICANT: Dillon, Davin C.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
FILE REFERENCE: 210121.534C1  
CURRENT APPLICATION NUMBER: US/09/822,827  
CURRENT FILING DATE: 2001-03-28  
NUMBER OF SEQ ID NOS: 982  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 113  
LENGTH: 553  
TYPE: PRT  
ORGANISM: Homo sapien  
US-09-822-827-113

Query Match 34.5%; Score 50; DB 10; Length 553;  
Best Local Similarity 45.0%; Pred. No. 9.3;  
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;  
QY 2 YSLPELDYFSATEPYISGQ 21  
DB 176 YLPAIDWTSALAPYLGTQ 195

RESULT 33  
US-09-115-453-113  
Sequence 113, Application US/09115453B  
Patent No. US20020090372A1  
GENERAL INFORMATION:  
APPLICANT: Xu, Jiangchun  
APPLICANT: Dillon, Davin C.  
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND  
FILE REFERENCE: 210121.427C4  
CURRENT APPLICATION NUMBER: US/09/115,453B  
CURRENT FILING DATE: 1998-07-14  
NUMBER OF SEQ ID NOS: 228  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 113  
LENGTH: 553  
TYPE: PRT  
ORGANISM: Homo sapien  
US-09-115-453-113

Query Match 34.5%; Score 50; DB 10; Length 553;  
Best Local Similarity 45.0%; Pred. No. 9.3;  
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;  
QY 2 YSLPELDYFSATEPYISGQ 21

Db 176 YLPAIDWTSALAPYLGTO 195

## RESULT 34

US-09-895-793-947  
; Sequence 947, Application US/09895793  
; Publication No. US20020192763A1  
; GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun  
APPLICANT: Dillon, Devin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Jiang, Yugu  
APPLICANT: Kaios, Michael D.  
APPLICANT: Reiter, Marc W.  
APPLICANT: Stolk, John A.  
APPLICANT: Day, Craig H.  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Carter, Darrick  
APPLICANT: Li, Samuel X.  
APPLICANT: Wang, Aljun  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Hedler, William T.  
APPLICANT: Henderson, Robert A.  
APPLICANT: Hurai, John  
APPLICANT: McNeill, Patricia D.  
APPLICANT: Houghton, Raymond L.  
APPLICANT: Vinals de Baesols, Carlota  
APPLICANT: Foy, Teresa  
APPLICANT: Finger, Gary R.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
FILE REFERENCE: 210121.534C2  
CURRENT APPLICATION NUMBER: US/09/895,793  
CURRENT FILING DATE: 2001-06-29  
NUMBER OF SEQ. ID NOS: 982  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 947  
LENGTH: 1079  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-895-793-947

Query Match 34.5%; Score 50; DB 9; Length 1079;  
Best Local Similarity 45.0%; Pred. No. 20;  
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

2 YSLPELDYFSATEPYISGQ 21  
Db 702 YLPAIDWTSALAPYLGTO 721

## RESULT 35

US-09-822-827-947  
; Sequence 947, Application US/09822827  
; Patent No. US20020081680A1  
; GENERAL INFORMATION:  
APPLICANT: Xu, Jiangchun  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
FILE REFERENCE: 210121.534C1  
CURRENT APPLICATION NUMBER: US/09/822,827  
CURRENT FILING DATE: 2001-03-28  
NUMBER OF SEQ. ID NOS: 982  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 947  
LENGTH: 1079  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-822-827-947

Query Match 34.5%; Score 50; DB 10; Length 1079;  
Best Local Similarity 45.0%; Pred. No. 20;

Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 2 YSLPELDYFSATEPYISGQ 21  
Db 702 YLPAIDWTSALAPYLGTO 721

## RESULT 36

US-09-972-484-11  
; Sequence 11, Application US/09972484  
; Patent No. US20020127245A1  
; GENERAL INFORMATION:  
APPLICANT: Miller, Timothy J.  
Klepfier, Sharon  
Reed, Albert Paul  
Jones, Elaine V.  
TITLE OF INVENTION: Canine Coronavirus S Gene and Uses  
Therefor  
NUMBER OF SEQUENCES: 59  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmtlKline Beecham Corporation - Corporate  
Patents  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406-2799  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/972,484  
FILING DATE: 05-Oct-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/494,151  
FILING DATE: 28-Jan-2000  
APPLICATION NUMBER: 08/331,625  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US 07/698,927  
FILING DATE: 13-MAY-1991  
APPLICATION NUMBER: US 07/613,066  
FILING DATE: 14-NOV-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Schreck, Patricia A.  
REGISTRATION NUMBER: 33,777  
REFERENCE/DOCKET NUMBER: SBC H85010-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 270-5015  
TELEFAX: (215) 270-5090  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 203 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 11:  
US-09-972-484-11

Query Match 31.4%; Score 45.5; DB 10; Length 203;  
Best Local Similarity 33.3%; Pred. No. 15;  
Matches 9; Conservative 11; Mismatches 6; Indels 1; Gaps 1;

QY 2 YSLPELDYF-SATEPYISQINEIXY 27  
Db 79 WTVPELPIDIFNATYLNLTGEINDLEF 105

RESULT 37  
US-09-972-484-52  
; Sequence 52, Application US/09972484

Patent No. US20020127245A1  
GENERAL INFORMATION:  
APPLICANT: Miller, Timothy J.  
Klepfisz, Sharon  
Reed, Albert Paul  
Jones, Elaine V.  
TITLE OF INVENTION: Canine Coronavirus S Gene and Uses  
Therefor  
NUMBER OF SEQUENCES: 59  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation - Corporate  
Patents  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406-2799  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/972,484  
FILING DATE: 05-Oct-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/494,151  
FILING DATE: 28-Jan-2000  
APPLICATION NUMBER: 08/331,625  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US 07/698,927  
FILING DATE: 13-MAY-1991  
APPLICATION NUMBER: US 07/613,066  
FILING DATE: 14-NOV-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Schreck, Patricia A.  
REGISTRATION NUMBER: 33,777  
REFERENCE/DOCKET NUMBER: SBC H85010-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 270-5015  
TELEFAX: (215) 270-5015  
INFORMATION FOR SEQ ID NO: 52:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1101 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 52:  
-09-972-484-52  
Query Match 31.4%; Score 45.5; DB 10; Length 1101;  
Best Local Similarity 33.3%; Pred. No. 98;  
Matches 9; Conservative 11; Mismatches 6; Indels 1; Gaps 1;  
QY 2 YSLPELDYE-PSATEPYISQGINEIXY 27  
Db 977 WTVPELPDIFNATYLNUTGEINDLEF 1003  
RESULT 38  
US-09-972-484-2  
Sequence 2, Application US/09972484  
Patent No. US20020127245A1  
GENERAL INFORMATION:  
APPLICANT: Miller, Timothy J.  
Klepfisz, Sharon  
Reed, Albert Paul  
Jones, Elaine V.  
TITLE OF INVENTION: Canine Coronavirus S Gene and Uses  
Therefor  
NUMBER OF SEQUENCES: 59  
CORRESPONDENCE ADDRESS:

ADDRESSEE: SmithKline Beecham Corporation - Corporate  
Patents  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406-2799  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/972,484  
FILING DATE: 05-Oct-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/494,151  
FILING DATE: 28-Jan-2000  
APPLICATION NUMBER: 08/331,625  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US 07/698,927  
FILING DATE: 13-MAY-1991  
APPLICATION NUMBER: US 07/613,066  
FILING DATE: 14-NOV-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Schreck, Patricia A.  
REGISTRATION NUMBER: 33,777  
REFERENCE/DOCKET NUMBER: SBC H85010-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 270-5015  
TELEFAX: (215) 270-5090  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1452 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-972-484-2  
Query Match 31.4%; Score 45.5; DB 10; Length 1452;  
Best Local Similarity 33.3%; Pred. No. 1.3e+02;  
Matches 9; Conservative 11; Mismatches 6; Indels 1; Gaps 1;  
QY 2 YSLPELDYE-PSATEPYISQGINEIXY 27  
Db 1328 WTVPELPDIFNATYLNUTGEINDLEF 1354  
RESULT 39  
US-09-832-355A-122  
Sequence 122, Application US/09832355A  
Publication No. US200302751A1  
GENERAL INFORMATION:  
APPLICANT: Kovsedl, Imre  
Klepfisz, Sharon  
Reed, Albert Paul  
Jones, Elaine V.  
TITLE OF INVENTION: VEGF FUSION PROTEINS  
FILE REFERENCE: 205654  
CURRENT APPLICATION NUMBER: US/09/832,355A  
CURRENT FILING DATE: 2001-04-10  
NUMBER OF SEQ ID NOS: 126  
SOFTWARE: Patentin version 3.0  
SEQ ID NO 122  
LENGTH: 280  
TYPE: PRT  
ORGANISM: Artificial sequence  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: ( )-( )  
OTHER INFORMATION: Synthetic  
US-09-832-355A-122

Query Match	31.0%	Score 45	DB 9	Length 280
Best Local Similarity	38.1%	Pred. No. 25		
Matches	8	Conservative	6	Mismatches 7
				Indels 0
				Gaps 0
QY	2	YSLPELDYRFATPEPIISQI	22	
		..   ..	..   ..	:
Db	45	FLPELDNCRSSSPFVSNV	65	

```

RESULT 40
US-09-832-355A-96
; Sequence 96, Application US/09832355A
; Publication No. US20030027751A1
; GENERAL INFORMATION:
; APPLICANT: Kovesdi, Imre
; APPLICANT: Kessler, Paul
; TITLE OF INVENTION: VEGF FUSION PROTEINS
; FILE REFERENCE: 205654
; CURRENT APPLICATION NUMBER: US/09/832,355A
; CURRENT FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 96
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1..1)
; OTHER INFORMATION: Synthetic
; NAME/KEY: misc.feature
; LOCATION: (2691..1272)
; OTHER INFORMATION: "Xaa" may be any amino acid
US-09-832-355A-96

Query Match          31.0%; Score 45; DB 9; Length 360;
Best Local Similarity 38.1%; Pred. No. 34;
Matches 8; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 2 YSLPELDYEFSAPEYISQI 22
   :|::|:|::|::|::|::|:
DB 45 FLPPENDCRSSSPFVSNV 65

```

Search completed: April 9, 2003, 14:12:03  
Job time : 8.33027 secs





GenCore version 5.1.4 p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 9, 2003, 14:17:20 ; Search time 37.945 Seconds  
(without alignments)  
302.901 Million cell updates/sec

Title: US-09-987-190-6

Perfect score: 188  
Sequence: 1 ATEPPTGQWNEIHYTKHQ.....YFKNLWVHWKAEERRFEF 188

Scoring table: OLIGO  
Gapco 60.0 , Gapext 60.0

Database: 248812 seqs, 61136040 residues

Word size : 0

Total number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database : Published Applications AA.\*

1: /cgn2\_6/prodata/2/pubppaa/US08\_NEW\_PUB pep.\*  
2: /cgn2\_6/prodata/2/pubppaa/PCT\_NEW\_PUB pep.\*  
3: /cgn2\_6/prodata/2/pubppaa/US06\_NEW\_PUB pep.\*  
4: /cgn2\_6/prodata/2/pubppaa/US06\_PUBCOMB pep.\*  
5: /cgn2\_6/prodata/2/pubppaa/US07\_NEW\_PUB pep.\*  
6: /cgn2\_6/prodata/2/pubppaa/US07\_PUBCOMB pep.\*  
7: /cgn2\_6/prodata/2/pubppaa/PCTUS\_PUBCOMB pep.\*  
8: /cgn2\_6/prodata/2/pubppaa/US08\_NEW\_PUB pep.\*  
9: /cgn2\_6/prodata/2/pubppaa/US08\_PUBCOMB pep.\*  
10: /cgn2\_6/prodata/2/pubppaa/US09\_NEW\_PUB pep.\*  
11: /cgn2\_6/prodata/2/pubppaa/US10\_NEW\_PUB pep.\*  
12: /cgn2\_6/prodata/2/pubppaa/US10\_PUBCOMB pep.\*  
13: /cgn2\_6/prodata/2/pubppaa/US60\_NEW\_PUB pep.\*  
14: /cgn2\_6/prodata/2/pubppaa/US60\_PUBCOMB pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	188	100.0	188	US-09-987-190-6	Sequence 6, Appl1
2	9	4.8	222	US-09-727-855B-5	Sequence 721, Appl
3	7	3.7	121	US-10-101-464A-721	Sequence 721, Appl
4	7	3.7	198	US-09-727-855B-7	Sequence 702, Appl
5	7	3.7	238	US-10-101-464A-702	Sequence 702, Appl
6	7	3.7	329	US-09-712-363-271	Sequence 271, Appl
7	7	3.7	506	US-09-815-242-11129	Sequence 11129, A
8	7	3.7	855	US-09-817-913-13	Sequence 13, Appl
9	7	3.7	855	US-09-817-538-13	Sequence 13, Appl
10	7	3.7	1333	US-09-738-626-4048	Sequence 4048, Ap
11	7	3.2	21	US-09-853-830-28	Sequence 28, Appl
12	6	3.2	30	US-09-987-190-2	Sequence 2, Appl1
13	6	3.2	39	US-09-864-761-36462	Sequence 36462, A
14	6	3.2	81	US-09-981-876-166	Sequence 166, App
15	6	3.2	84	US-09-148-545-166	Sequence 166, App
16	6	3.2	84	US-09-867-550-1282	Sequence 1282, Ap
17	6	3.2	95	US-09-933-533-12	Sequence 12, Appl
18	6	3.2	104	US-09-515-806-28	Sequence 28, Appl
19	6	3.2	123	US-09-738-626-4908	Sequence 4908, Ap

20	6	3.2	128	US-10-116-273-16	Sequence 16, Appl
21	6	3.2	129	US-10-116-273-15	Sequence 15, Appl
22	6	3.2	130	US-10-116-273-14	Sequence 14, Appl
23	6	3.2	131	US-10-116-273-13	Sequence 13, Appl
24	6	3.2	132	US-10-116-273-12	Sequence 12, Appl
25	6	3.2	133	US-10-116-273-11	Sequence 11, Appl
26	6	3.2	134	US-10-116-273-10	Sequence 10, Appl
27	6	3.2	135	US-10-116-273-9	Sequence 9, Appl1
28	6	3.2	136	US-10-116-273-8	Sequence 8, Appl1
29	6	3.2	137	US-10-116-273-7	Sequence 7, Appl1
30	6	3.2	138	US-10-116-273-6	Sequence 6, Appl1
31	6	3.2	139	US-10-116-273-5	Sequence 5, Appl1
32	6	3.2	140	US-10-116-273-4	Sequence 4, Appl1
33	6	3.2	141	US-10-116-273-3	Sequence 3, Appl1
34	6	3.2	142	US-10-116-273-2	Sequence 2, Appl1
35	6	3.2	143	US-09-308-207-60	Sequence 60, Appl
36	6	3.2	144	US-10-116-273-1	Sequence 1, Appl1
37	6	3.2	151	US-10-147-874-6	Sequence 6, Appl1
38	6	3.2	153	US-09-765-272-224	Sequence 224, App
39	6	3.2	160	US-09-864-761-48707	Sequence 48707, A
40	6	3.2	161	US-10-101-464A-548	Sequence 548, App
41	6	3.2	168	US-09-746-801A-63	Sequence 63, Appl
42	6	3.2	197	US-09-738-626-6727	Sequence 11, Appl
43	6	3.2	200	US-09-738-626-6727	Sequence 6727, Ap
44	6	3.2	200	US-09-818-564-2	Sequence 2, Appl1
45	6	3.2	214	US-09-924-256A-74	Sequence 74, Appl
46	6	3.2	222	US-09-925-301-1228	Sequence 1228, Ap
47	6	3.2	229	US-09-939-980-519	Sequence 519, App
48	6	3.2	231	US-10-006-922-6	Sequence 6, Appl1
49	6	3.2	231	US-10-081-864-4	Sequence 4, Appl1
50	6	3.2	232	US-09-978-248-6	Sequence 6, Appl1
51	6	3.2	232	US-09-978-248-7	Sequence 7, Appl1
52	6	3.2	245	US-09-815-242-10921	Sequence 10921, A
53	6	3.2	246	US-09-815-242-15461	Sequence 5461, Ap
54	6	3.2	246	US-09-815-242-12123	Sequence 12123, A
55	6	3.2	246	US-09-815-242-12803	Sequence 12803, A
56	6	3.2	246	US-09-815-242-13100	Sequence 13100, A
57	6	3.2	247	US-09-912-787-2	Sequence 2, Appl1
58	6	3.2	247	US-09-912-787-55	Sequence 55, Appl
59	6	3.2	247	US-10-014-326-41	Sequence 41, Appl
60	6	3.2	247	US-10-014-326-66	Sequence 66, Appl
61	6	3.2	268	US-09-925-301-1044	Sequence 1044, Ap
62	6	3.2	286	US-09-866-050A-137	Sequence 137, App
63	6	3.2	301	US-09-925-301-1196	Sequence 1196, Ap
64	6	3.2	328	US-10-164-433-4	Sequence 4, Appl1
65	6	3.2	328	US-09-738-626-4022	Sequence 4022, Ap
66	6	3.2	337	US-09-815-242-5606	Sequence 5606, Ap
67	6	3.2	359	US-09-759-097-4	Sequence 4, Appl1
68	6	3.2	359	US-09-759-097-4	Sequence 12457, A
69	6	3.2	392	US-09-759-097-2	Sequence 2, Appl1
70	6	3.2	399	US-09-817-513A-2	Sequence 2, Appl1
71	6	3.2	422	US-09-815-242-11351	Sequence 11351, A
72	6	3.2	428	US-10-062-254-250	Sequence 250, App
73	6	3.2	432	US-10-062-254-242	Sequence 242, App
74	6	3.2	444	US-09-815-242-5785	Sequence 5785, Ap
75	6	3.2	448	US-09-738-626-4036	Sequence 4036, Ap
76	6	3.2	452	US-09-815-242-11781	Sequence 11781, A
77	6	3.2	455	US-09-955-999-110	Sequence 110, App
78	6	3.2	470	US-10-067-668-12	Sequence 12, Appl
79	6	3.2	470	US-10-043-487-10	Sequence 10, Appl
80	6	3.2	495	US-09-738-626-3801	Sequence 3801, Ap
81	6	3.2	496	US-10-067-668-10	Sequence 10, Appl
82	6	3.2	496	US-09-945-301-4	Sequence 4, Appl1
83	6	3.2	510	US-09-738-626-6798	Sequence 6798, Ap
84	6	3.2	511	US-09-733-882-11	Sequence 11, Appl
85	6	3.2	511	US-09-731-872-278	Sequence 278, App
86	6	3.2	541	US-09-731-872-330	Sequence 330, App
87	6	3.2	541	US-09-731-872-330	Sequence 300, App
88	6	3.2	574	US-09-992-598-340	Sequence 340, App
89	6	3.2	574	US-09-989-293A-340	Sequence 340, App
90	6	3.2	574	US-09-989-293-340	Sequence 340, App
91	6	3.2	574	US-09-990-444-340	Sequence 340, App
92	6	3.2	574	US-09-989-730-340	Sequence 340, App

93	3	2	574	9	US-09-990-436-340	Sequence 340, App	166	3	2	574	9	US-10-184-652-234	Sequence 234, App		
94	6	3	2	574	9	US-09-991-181-340	Sequence 340, App	167	6	3	2	574	9	US-10-187-594-234	Sequence 234, App
95	6	3	2	574	9	US-09-993-687-340	Sequence 340, App	168	6	3	2	574	9	US-10-187-556-234	Sequence 234, App
96	6	3	2	574	9	US-09-989-734-340	Sequence 340, App	169	6	3	2	574	9	US-10-187-745-234	Sequence 234, App
97	6	3	2	574	9	US-09-997-653-340	Sequence 234, App	170	6	3	2	574	9	US-10-187-885-234	Sequence 234, App
98	6	3	2	574	9	US-10-174-590-234	Sequence 234, App	171	6	3	2	574	9	US-10-187-886-234	Sequence 234, App
99	6	3	2	574	9	US-10-176-758-234	Sequence 234, App	172	6	3	2	574	9	US-10-199-464-234	Sequence 234, App
100	6	3	2	574	9	US-10-175-737-234	Sequence 234, App	173	6	3	2	574	9	US-10-176-751-234	Sequence 234, App
101	6	3	2	574	9	US-09-993-667-340	Sequence 234, App	174	6	3	2	574	9	US-10-176-760-234	Sequence 234, App
102	6	3	2	574	9	US-10-173-706-234	Sequence 234, App	175	6	3	2	574	9	US-10-176-990-234	Sequence 234, App
103	6	3	2	574	9	US-10-175-738-234	Sequence 234, App	176	6	3	2	574	9	US-10-180-541-234	Sequence 234, App
104	6	3	2	574	9	US-10-175-752-234	Sequence 234, App	177	6	3	2	574	9	US-10-180-542-234	Sequence 234, App
105	6	3	2	574	9	US-10-176-482-234	Sequence 234, App	178	6	3	2	574	9	US-10-180-548-234	Sequence 234, App
106	6	3	2	574	9	US-10-176-757-234	Sequence 234, App	179	6	3	2	574	9	US-10-180-551-234	Sequence 234, App
107	6	3	2	574	9	US-10-176-913-234	Sequence 234, App	180	6	3	2	574	9	US-10-180-998-234	Sequence 234, App
108	6	3	2	574	9	US-10-180-552-234	Sequence 234, App	181	6	3	2	574	9	US-10-180-999-234	Sequence 234, App
109	6	3	2	574	9	US-10-180-557-234	Sequence 234, App	182	6	3	2	574	9	US-10-183-013-234	Sequence 234, App
110	6	3	2	574	9	US-09-990-438-340	Sequence 340, App	183	6	3	2	574	9	US-10-184-612-234	Sequence 234, App
111	6	3	2	574	9	US-09-990-562-340	Sequence 340, App	184	6	3	2	574	9	US-10-184-630-234	Sequence 234, App
112	6	3	2	574	9	US-09-997-428-340	Sequence 340, App	185	6	3	2	574	9	US-10-184-616-234	Sequence 234, App
113	6	3	2	574	9	US-09-997-666-340	Sequence 340, App	186	6	3	2	574	9	US-10-184-622-234	Sequence 234, App
114	6	3	2	574	9	US-10-173-700-234	Sequence 234, App	187	6	3	2	574	9	US-10-184-628-234	Sequence 234, App
115	6	3	2	574	9	US-10-174-572-234	Sequence 234, App	188	6	3	2	574	9	US-10-184-629-234	Sequence 234, App
116	6	3	2	574	9	US-10-174-579-234	Sequence 234, App	189	6	3	2	574	9	US-10-184-651-234	Sequence 234, App
117	6	3	2	574	9	US-10-174-582-234	Sequence 234, App	190	6	3	2	574	9	US-10-184-631-234	Sequence 234, App
118	6	3	2	574	9	US-10-174-588-234	Sequence 234, App	191	6	3	2	574	9	US-10-184-632-234	Sequence 234, App
119	6	3	2	574	9	US-10-175-739-234	Sequence 234, App	192	6	3	2	574	9	US-10-184-636-234	Sequence 234, App
120	6	3	2	574	9	US-10-175-740-234									

223	6	3.2	574	9	US-10-184-626-234	Sequence 234, App	312	6	3.2	574	9	US-10-202-469-234	Sequence 234, App
229	6	3.2	574	9	US-10-184-627-234	Sequence 234, App	313	6	3.2	574	9	US-10-202-470-234	Sequence 234, App
240	6	3.2	574	9	US-10-184-645-234	Sequence 234, App	314	6	3.2	574	9	US-10-202-476-234	Sequence 234, App
241	6	3.2	574	9	US-10-184-654-234	Sequence 234, App	315	6	3.2	574	9	US-10-202-934-234	Sequence 234, App
242	6	3.2	574	9	US-10-184-655-234	Sequence 234, App	316	6	3.2	574	9	US-10-202-935-234	Sequence 234, App
243	6	3.2	574	9	US-10-188-775-234	Sequence 234, App	317	6	3.2	574	9	US-10-202-936-234	Sequence 234, App
244	6	3.2	574	9	US-10-188-776-234	Sequence 234, App	318	6	3.2	574	9	US-10-202-939-234	Sequence 234, App
245	6	3.2	574	9	US-10-194-462-234	Sequence 234, App	319	6	3.2	574	9	US-10-205-504-234	Sequence 234, App
246	6	3.2	574	9	US-10-195-902-234	Sequence 234, App	320	6	3.2	574	9	US-10-205-509-234	Sequence 234, App
247	6	3.2	574	9	US-10-196-743-234	Sequence 234, App	321	6	3.2	574	9	US-10-205-895-234	Sequence 234, App
248	6	3.2	574	9	US-10-196-745-234	Sequence 234, App	322	6	3.2	574	9	US-10-205-899-234	Sequence 234, App
249	6	3.2	574	9	US-10-196-760-234	Sequence 234, App	323	6	3.2	574	9	US-10-205-900-234	Sequence 234, App
250	6	3.2	574	9	US-10-196-762-234	Sequence 234, App	324	6	3.2	574	9	US-10-205-909-234	Sequence 234, App
251	6	3.2	574	9	US-10-197-695-234	Sequence 234, App	325	6	3.2	574	9	US-09-990-443-340	Sequence 340, App
252	6	3.2	574	9	US-09-990-437-340	Sequence 340, App	326	6	3.2	574	9	US-09-990-726-340	Sequence 340, App
253	6	3.2	574	9	US-10-179-506-234	Sequence 234, App	327	6	3.2	574	9	US-09-997-559-340	Sequence 340, App
254	6	3.2	574	9	US-10-176-484-234	Sequence 234, App	328	6	3.2	574	9	US-09-997-601-340	Sequence 340, App
255	6	3.2	574	9	US-10-176-753-234	Sequence 234, App	329	6	3.2	574	9	US-10-183-002-234	Sequence 234, App
256	6	3.2	574	9	US-10-176-917-234	Sequence 234, App	330	6	3.2	574	9	US-10-184-621-234	Sequence 234, App
257	6	3.2	574	9	US-10-176-982-234	Sequence 234, App	331	6	3.2	574	9	US-10-184-638-234	Sequence 234, App
258	6	3.2	574	9	US-10-179-506-234	Sequence 234, App	332	6	3.2	574	9	US-10-187-752-234	Sequence 234, App
259	6	3.2	574	9	US-10-179-513-234	Sequence 234, App	333	6	3.2	574	9	US-10-187-887-234	Sequence 234, App
260	6	3.2	574	9	US-10-179-514-234	Sequence 234, App	334	6	3.2	574	9	US-10-194-461-234	Sequence 234, App
261	6	3.2	574	9	US-10-179-522-234	Sequence 234, App	335	6	3.2	574	9	US-10-195-882-234	Sequence 234, App
262	6	3.2	574	9	US-10-180-556-234	Sequence 234, App	336	6	3.2	574	9	US-10-196-751-234	Sequence 234, App
263	6	3.2	574	9	US-10-180-560-234	Sequence 234, App	337	6	3.2	574	9	US-10-197-694-234	Sequence 234, App
264	6	3.2	574	9	US-10-183-015-234	Sequence 234, App	338	6	3.2	574	9	US-10-197-697-234	Sequence 234, App
265	6	3.2	574	9	US-10-184-615-234	Sequence 234, App	339	6	3.2	574	9	US-10-197-707-234	Sequence 234, App
266	6	3.2	574	9	US-10-184-620-234	Sequence 234, App	340	6	3.2	574	9	US-10-199-	

385	6	3.2	574	12	US-10-052-586-234	Sequence 234, App	458	5	2.7	20	10	US-09-215-655-45	Sequence 45, App1
386	6	3.2	578	10	US-09-823-356-10	Sequence 10, App1	459	5	2.7	20	10	US-09-922-261-381	Sequence 381, App
387	6	3.2	580	10	US-09-866-050A-307	Sequence 307, App	460	5	2.7	21	10	US-09-853-830-51	Sequence 51, App1
388	6	3.2	608	10	US-09-767-041-39	Sequence 39, App1	461	5	2.7	21	12	US-10-153-034-119	Sequence 119, App
389	6	3.2	734	10	US-09-684-998-26	Sequence 26, App1	462	5	2.7	22	9	US-09-759-130B-458	Sequence 458, App
390	6	3.2	757	10	US-09-928-175-24	Sequence 24, App1	463	5	2.7	23	9	US-10-178-043-6	Sequence 6, App11
391	6	3.2	785	9	US-09-738-626-5445	Sequence 5445, Ap	464	5	2.7	23	10	US-09-734-41-5	Sequence 5, App1
392	6	3.2	807	9	US-10-029-217A-2	Sequence 2, App11	465	5	2.7	24	10	US-10-036-729-5	Sequence 5, App1
393	6	3.2	807	9	US-10-029-217A-26	Sequence 26, App1	466	5	2.7	24	10	US-09-922-261-244	Sequence 244, App
394	6	3.2	811	9	US-09-975-719-93	Sequence 93, App1	467	5	2.7	25	10	US-09-739-907-11	Sequence 118, App
395	6	3.2	833	10	US-09-691-216-3	Sequence 3, App11	468	5	2.7	25	10	US-09-922-261-379	Sequence 379, App
396	6	3.2	838	9	US-09-712-363-150	Sequence 150, App	469	5	2.7	26	10	US-09-764-877-184	Sequence 1384, Ap
397	6	3.2	838	9	US-10-219-219-4	Sequence 4, App11	470	5	2.7	27	9	US-09-747-802-2	Sequence 2, App1
398	6	3.2	855	9	US-10-072-094-9	Sequence 9, App11	471	5	2.7	28	10	US-09-864-761-41156	Sequence 47056, A
399	6	3.2	864	10	US-09-883-096-2	Sequence 2, App11	472	5	2.7	28	10	US-09-864-761-48026	Sequence 48026, A
400	6	3.2	875	9	US-10-115-515-10	Sequence 10, App1	473	5	2.7	29	10	US-09-864-761-33337	Sequence 34337, A
401	6	3.2	875	9	US-10-115-515-23	Sequence 23, App1	474	5	2.7	29	10	US-09-922-261-377	Sequence 377, App
402	6	3.2	876	9	US-09-738-626-4789	Sequence 4789, Ap	475	5	2.7	29	10	US-09-922-261-435	Sequence 435, App
403	6	3.2	901	10	US-09-815-242-10034	Sequence 10034, A	476	5	2.7	30	10	US-09-864-761-43714	Sequence 43714, A
404	6	3.2	901	10	US-10-029-217A-31	Sequence 31, App1	477	5	2.7	30	12	US-10-038-045-9	Sequence 9, App11
405	6	3.2	935	9	US-10-029-217A-32	Sequence 32, App1	478	5	2.7	31	10	US-09-823-901-8	Sequence 8, App11
406	6	3.2	935	9	US-10-029-217A-28	Sequence 28, App1	479	5	2.7	31	10	US-09-864-761-45334	Sequence 45334, A
407	6	3.2	938	9	US-10-029-217A-29	Sequence 29, App1	480	5	2.7	32	10	US-09-864-761-34360	Sequence 34360, A
408	6	3.2	938	9	US-10-078-531-3	Sequence 3, App11	481	5	2.7	32	10	US-09-142-755-14	Sequence 14, App1
409	6	3.2	951	9	US-09-884-696-3	Sequence 3, App11	482	5	2.7	33	9	US-09-852-424-113	Sequence 113, App
410	6	3.2	953	9	US-10-078-531-5	Sequence 5, App11	483	5	2.7	33	9	US-09-852-424-114	Sequence 114, App
411	6	3.2	963	9	US-10-078-531-8	Sequence 8, App11	484	5	2.7	33	9	US-09-852-424-115	Sequence 115, App
412	6	3.2	969	9	US-10-078-531-4	Sequence 4, App11	485	5	2.7	33	9	US-09-835-107-18	Sequence 18, App1
413	6	3.2	970	9	US-10-078-531-6	Sequence 6, App11	486	5	2.7	33	9	US-09-835-107-16	Sequence 16, App1
414	6	3.2	971	9	US-10-078-531-7	Sequence 7, App11	487	5	2.7	33	9	US-10-081-816-78	Sequence 78, App1
415	6	3.2	1008	9	US-09-712-363-287	Sequence 287, App	488	5	2.7	33	10	US-09-864-761-48221	Sequence 48221, A
416	6	3.2	1094	10	US-09-815-242-4930	Sequence 4930, Ap	490	5	2.7	34	9	US-09-852-424-75	Sequence 75, App1
417	6	3.2	1109	10	US-09-815-242-10648	Sequence 10648, A	491	5	2.7	34	9	US-09-852-424-82	Sequence 82, App1
418	6	3.2	1109	10	US-09-815-242-10648	Sequence 20, App1	492	5	2.7	34	9	US-09-852-424-83	Sequence 83, App1
419	6	3.2	1161	8	US-09-866-582-36	Sequence 36, App1	493	5	2.7	34	9	US-09-852-424-84	Sequence 84, App1
420	6	3.2	1198	10	US-10-001-215-5	Sequence 5, App11	494	5	2.7	34	9	US-09-852-424-85	Sequence 85, App1
421	6	3.2	1241	12	US-09-801-368-364	Sequence 364, App	495	5	2.7	34	9	US-09-852-424-98	Sequence 98, App1
422	6	3.2	1247	10	US-09-801-368-364	Sequence 979, App	496	5	2.7	34	9	US-09-852-424-100	Sequence 100, App
423	6	3.2	1250	10	US-09-801-368-364	Sequence 1071, A	497	5	2.7	34	9	US-09-852-424-123	Sequence 123, App
424	6	3.2	1270	9	US-09-815-242-10717	Sequence 5118, Ap	498	5	2.7	34	9	US-09-852-424-125	Sequence 125, App
425	6	3.2	1271	10	US-09-951-402-3	Sequence 3, App11	499	5	2.7	34	9	US-09-835-107-19	Sequence 19, App1
426	6	3.2	1317	10	US-09-951-402-3	Sequence 3, App11	500	5	2.7	34	9	US-09-835-107-19	Sequence 19, App1
427	6	3.2	2037	10	US-09-801-268-172	Sequence 172, App	501	5	2.7	34	9	US-10-013-939-3	Sequence 3, App11
428	6	3.2	2037	10	US-09-801-268-172	Sequence 28, App1	502	5	2.7	35	9	US-10-050-704-193	Sequence 193, App
429	6	3.2	2037	10	US-09-801-268-172	Sequence 37, App1	503	5	2.7	35	9	US-09-864-761-40859	Sequence 40859, A
430	6	3.2	3092	10	US-09-801-268-172	Sequence 17, App1	504	5	2.7	35	10	US-09-864-761-46218	Sequence 46218, A
431	6	3.2	11	10	US-10-036-729-17	Sequence 18, App1	505	5	2.7	35	10	US-09-864-761-47179	Sequence 47179, A
432	6	3.2	12	9	US-10-036-729-18	Sequence 19, App1	506	5	2.7	37	9	US-10-097-065-360	Sequence 360, App
433	6	3.2	12	9	US-10-036-729-19	Sequence 20, App1	507	5	2.7	37	9	US-09-798-889-83	Sequence 889-83
434	6	3.2	12	9	US-10-036-729-20	Sequence 180, App	508	5	2.7	38	9	US-09-864-480-418	Sequence 418, App
435	6	3.2	13	9	US-09-776-724A-180	Sequence 12, App1	509	5	2.7	39	10	US-09-864-761-41311	Sequence 41311, A
436	6	3.2	14	9	US-10-024-133A-45	Sequence 45, App1	510	5	2.7	39	10	US-09-864-761-43920	Sequence 43920, A
437	6	3.2	14	10	US-09-791-378-328	Sequence 328, App	511	5	2.7	42	9	US-10-012-542-337	Sequence 337, App
438	6	3.2	14	10	US-09-791-378-328	Sequence 390, App	512	5	2.7	43	9	US-09-913-153-122	Sequence 2, App11
439	6	3.2	14	10	US-09-791-378-328	Sequence 466, App	513	5	2.7	43	10	US-09-864-761-43985	Sequence 43985, A
440	6	3.2	14	10	US-09-791-378-328	Sequence 17, App1	514	5	2.7	43	10	US-10-012-542-228	Sequence 228, App
441	6	3.2	15	9	US-09-826-290-197	Sequence 197, App	515	5	2.7	45	9	US-09-864-761-44415	Sequence 44415, A
442	6	3.2	15	9	US-09-826-290-197	Sequence 287, App	516	5	2.7	46	9	US-09-801-936-37	Sequence 37, App1
443	6	3.2	15	9	US-09-826-290-197	Sequence 290, App	517	5	2.7	46	9	US-09-813-153-122	Sequence 122, App
444	6	3.2	15	9	US-09-826-290-197	Sequence 82, App1	518	5	2.7	46	10	US-09-864-761-3082	Sequence 3082, A
445	6	3.2	15	10	US-09-923-684-14	Sequence 14, App1	519	5	2.7	46	10	US-09-864-761-42828	Sequence 42828, A
446	6	3.2	15	10	US-09-923-684-14	Sequence 415, App	520	5	2.7	47	10	US-09-864-761-42971	Sequence 42971, A
447	6	3.2	16	9	US-09-791-378-415	Sequence 117, App	521	5	2.7	47	10	US-09-925-301-1636	Sequence 1636, Ap
448	6	3.2	16	9	US-09-791-378-415	Sequence 4, App11	522	5	2.7	49	9	US-09-764-872-309	Sequence 309, App
449	6	3.2	19	10	US-10-045-180A-4	Sequence 21, App1	523	5	2.7	50	10	US-09-864-761-43011	Sequence 43011, A
450	6	3.2	19	10	US-09-193-538-21	Sequence 40545, A	524	5	2.7	51	9	US-10-092-154-646	Sequence 646, App
451	6	3.2	19	10	US-09-864-761-40545	Sequence 85, App1	525	5	2.7	51	10	US-09-864-761-43411	Sequence 43411, A
452	6	3.2	19	10	US-09-864-761-40545	Sequence 4, App1	526	5	2.7	51	10	US-09-764-847-646	Sequence 646, App
453	6	3.2	19	10	US-09-864-761-40545	Sequence 4, App1	527	5	2.7	51	10	US-09-764-847-646	Sequence 646, App
454	6	3.2	19	10	US-09-864-761-40545	Sequence 4, App1	528	5	2.7	51	10	US-09-764-847-646	Sequence 646, App
455	6	3.2	19	10	US-09-864-761-40545	Sequence 4, App1	529	5	2.7	51	10	US-09-764-847-646	Sequence 646, App
456	6	3.2	19	10	US-09-864-761-40545	Sequence 4, App1	530	5	2.7	51	10	US-09-764-847-646	Sequence 646, App
457	6	3.2	19	10	US-09-864-761-40545	Sequence 4, App1	530	5	2.7	51	10	US-09-764-847-646	Sequence 646, App

531	5	2.7	51	12	US-10-067-422-25	Sequence 25, Appl	604	5	2.7	84	10	US-09-925-299-1207	Sequence 1207, Ap
532	5	2.7	52	10	US-09-843-845-14	Sequence 14, Appl	605	5	2.7	85	10	US-09-861-400-9	Sequence 9, Appl1
533	5	2.7	53	10	US-09-864-761-40055	Sequence 40055, A	606	5	2.7	86	10	US-09-864-761-38602	Sequence 38602, A
534	5	2.7	53	10	US-09-864-761-40784	Sequence 40784, A	607	5	2.7	85	10	US-09-864-761-48904	Sequence 48904, A
535	5	2.7	53	10	US-09-864-761-43750	Sequence 43750, A	608	5	2.7	86	10	US-09-864-761-47312	Sequence 47312, A
536	5	2.7	54	9	US-09-925-299-1361	Sequence 1361, Ap	609	5	2.7	86	10	US-09-987-108-21	Sequence 21, Appl
537	5	2.7	54	10	US-09-925-299-1361	Sequence 1361, Ap	610	5	2.7	88	9	US-10-101-4644-78	Sequence 78, Appl
538	5	2.7	55	10	US-09-864-761-36786	Sequence 36786, A	611	5	2.7	89	9	US-10-101-4644-82	Sequence 82, Appl
539	5	2.7	56	10	US-09-867-550-178	Sequence 178, App	612	5	2.7	89	9	US-10-102-627-47	Sequence 47, Appl
540	5	2.7	57	10	US-09-864-761-37163	Sequence 37163, A	613	5	2.7	90	9	US-10-067-514-12	Sequence 12, Appl
541	5	2.7	58	9	US-09-925-299-1300	Sequence 1300, Ap	614	5	2.7	90	10	US-09-985-911-25	Sequence 25, Appl
542	5	2.7	58	10	US-09-864-761-47890	Sequence 47890, A	615	5	2.7	90	10	US-09-985-911-25	Sequence 25, Appl
543	5	2.7	58	10	US-09-925-299-1300	Sequence 1300, Ap	616	5	2.7	91	9	US-10-091-504-874	Sequence 874, Appl
544	5	2.7	59	9	US-09-796-692-745	Sequence 745, App	617	5	2.7	91	10	US-09-864-761-33756	Sequence 33756, A
545	5	2.7	59	10	US-09-864-761-44105	Sequence 44105, A	618	5	2.7	91	10	US-09-764-869-874	Sequence 874, App
546	5	2.7	59	10	US-09-864-761-45268	Sequence 45268, A	619	5	2.7	92	10	US-09-864-761-1507	Sequence 1507, Ap
547	5	2.7	60	12	US-10-060-509-17	Sequence 17, Appl	620	5	2.7	92	9	US-10-063-559-2	Sequence 2, Appl1
548	5	2.7	60	12	US-10-060-509-21	Sequence 21, Appl	621	5	2.7	92	9	US-10-091-504-693	Sequence 693, App
549	5	2.7	60	12	US-10-060-506-17	Sequence 17, Appl	622	5	2.7	92	10	US-09-864-761-36307	Sequence 36307, A
550	5	2.7	60	12	US-10-060-506-21	Sequence 21, Appl	623	5	2.7	92	10	US-09-764-869-693	Sequence 693, App
551	5	2.7	61	9	US-09-925-299-1331	Sequence 1331, App	624	5	2.7	92	10	US-09-764-877-1507	Sequence 1507, Ap
552	5	2.7	61	10	US-09-864-761-37344	Sequence 37344, A	625	5	2.7	93	9	US-09-738-626-6630	Sequence 6630, Ap
553	5	2.7	61	10	US-09-925-299-1331	Sequence 1331, Ap	626	5	2.7	93	9	US-09-893-519A-68	Sequence 68, Appl
554	5	2.7	62	9	US-09-984-245-240	Sequence 240, App	627	5	2.7	94	9	US-10-045-180A-3	Sequence 3, Appl1
555	5	2.7	62	9	US-09-966-662-240	Sequence 240, App	628	5	2.7	94	9	US-10-045-180A-10	Sequence 10, Appl
556	5	2.7	62	9	US-09-983-966-240	Sequence 240, App	629	5	2.7	94	9	US-10-045-180A-12	Sequence 12, Appl
557	5	2.7	62	9	US-10-091-504-1065	Sequence 1065, Ap	630	5	2.7	94	9	US-09-764-872-312	Sequence 312, App
558	5	2.7	62	10	US-09-764-869-1065	Sequence 1065, Ap	631	5	2.7	94	12	US-10-013-770-3	Sequence 3, Appl1
559	5	2.7	63	9	US-09-984-245-320	Sequence 320, App	632	5	2.7	95	10	US-09-764-887-269	Sequence 269, App
560	5	2.7	63	9	US-09-966-662-320	Sequence 320, App	633	5	2.7	95	10	US-09-917-340-76	Sequence 76, Appl
561	5	2.7	63	9	US-09-983-966-320	Sequence 320, App	634	5	2.7	96	10	US-09-800-729-119	Sequence 119, App
562	5	2.7	64	10	US-09-864-761-39808	Sequence 39808, A	635	5	2.7	96	10	US-09-917-340-89	Sequence 89, Appl
563	5	2.7	64	10	US-09-925-297-652	Sequence 652, App	636	5	2.7	96	10	US-09-917-340-90	Sequence 90, Appl
564	5	2.7	65	9	US-09-984-245-241	Sequence 241, App	637	5	2.7	97	9	US-10-045-180A-9	Sequence 9, Appl1
565	5	2.7	65	9	US-09-966-662-241	Sequence 241, App	638	5	2.7	97	10	US-09-864-761-33563	Sequence 33563, A
566	5	2.7	65	9	US-09-983-966-241	Sequence 241, App	639	5	2.7	97	10	US-09-864-761-40162	Sequence 40162, A
567	5	2.7	65	10	US-09-864-761-34093	Sequence 34093, A	640	5	2.7	97	10	US-09-864-761-41860	Sequence 41860, A
568	5	2.7	65	10	US-09-864-761-34172	Sequence 34172, A	641	5	2.7	97	10	US-09-884-441-482	Sequence 482, App
569	5	2.7	65	10	US-09-864-761-42072	Sequence 42072, A	642	5	2.7	98	10	US-09-815-242-13199	Sequence 13199, A
570	5	2.7	66	9	US-09-974-879-183	Sequence 183, App	643	5	2.7	99	1	US-08-781-986A-5523	Sequence 5223, Ap
571	5	2.7	66	10	US-09-861-400-10	Sequence 10, Appl	644	5	2.7	99	9	US-09-796-692-1588	Sequence 1588, Ap
572	5	2.7	67	9	US-09-738-626-3695	Sequence 3695, Appl	645	5	2.7	100	9	US-10-045-180A-11	Sequence 11, Appl
573	5	2.7	67	9	US-10-092-154-958	Sequence 958, App	646	5	2.7	100	10	US-09-864-761-35054	Sequence 35054, A
574	5	2.7	67	10	US-09-864-761-40129	Sequence 40129, A	647	5	2.7	100	10	US-09-840-459-37	Sequence 37, Appl
575	5	2.7	67	10	US-09-764-898-196	Sequence 196, App	648	5	2.7	101	9	US-09-948-820-59	Sequence 59, Appl
576	5	2.7	67	10	US-09-764-847-958	Sequence 958, App	649	5	2.7	101	9	US-10-083-357-714	Sequence 714, App
577	5	2.7	69	10	US-09-864-761-36291	Sequence 36291, A	650	5	2.7	101	10	US-09-815-242-5166	Sequence 5166, Ap
578	5	2.7	70	9	US-09-796-692-2365	Sequence 2365, Ap	651	5	2.7	101	10	US-09-764-877-1555	Sequence 1255, Ap
579	5	2.7	71	10	US-09-864-761-40846	Sequence 40846, A	652	5	2.7	101	10	US-09-840-459-42	Sequence 42, Appl
580	5	2.7	71	10	US-09-867-550-844	Sequence 844, App	653	5	2.7	101	10	US-09-840-459-46	Sequence 46, Appl
581	5	2.7	72	10	US-09-925-297-724	Sequence 724, App	654	5	2.7	102	9	US-10-004-381-20	Sequence 20, Appl
582	5	2.7	73	9	US-09-796-692-768	Sequence 768, App	655	5	2.7	104	10	US-09-939-980-255	Sequence 265, App
583	5	2.7	74	10	US-09-110-716-35	Sequence 35, Appl	656	5	2.7	105	9	US-09-808-602-2	Sequence 2, Appl1
584	5	2.7	74	10	US-09-864-761-41690	Sequence 41690, A	657	5	2.7	106	9	US-10-138-516-4	Sequence 4, Appl1
585	5	2.7	74	10	US-09-867-550-1832	Sequence 1832, Ap	658	5	2.7	106	9	US-10-146-130-6	Sequence 6, Appl1
586	5	2.7	76	9	US-09-749-637A-339	Sequence 339, App	659	5	2.7	106	9	US-10-092-934-5	Sequence 5, Appl1
587	5	2.7	76	10	US-09-864-761-41132	Sequence 41132, A	660	5	2.7	108	10	US-09-899-896-9	Sequence 9, Appl1
588	5	2.7	77	9	US-09-749-637A-59	Sequence 59, Appl	661	5	2.7	109	9	US-09-965-528-8	Sequence 8, Appl1
589	5	2.7	77	9	US-09-749-637A-77	Sequence 77, Appl	662	5	2.7	109	9	US-09-796-692-2460	Sequence 2460, Ap
590	5	2.7	79	10	US-09-922-261-433	Sequence 433, App	663	5	2.7	109	10	US-09-924-946-6	Sequence 6, Appl
591	5	2.7	79	10	US-09-884-441-460	Sequence 460, App	664	5	2.7	111	9	US-10-101-464A-722	Sequence 722, App
592	5	2.7	80	10	US-09-864-761-46010	Sequence 46010, A	665	5	2.7	111	10	US-09-867-550-258	Sequence 258, App
593	5	2.7	80	10	US-09-864-761-47469	Sequence 47469, A	666	5	2.7	111	10	US-09-934-054-5	Sequence 5, Appl1
594	5	2.7	80	10	US-09-862-179A-23	Sequence 23, Appl	667	5	2.7	112	10	US-09-978-244-14	Sequence 14, Appl
595	5	2.7	81	8	US-08-424-550B-187	Sequence 187, App	668	5	2.7	112	10	US-09-921-397-29	Sequence 29, Appl
596	5	2.7	81	9	US-10-125-540-515	Sequence 515, App	669	5	2.7	113	10	US-09-925-300-1011	Sequence 1011, Ap
597	5	2.7	81	10	US-09-764-870-515	Sequence 515, App	670	5	2.7	114	9	US-09-759-130B-198	Sequence 198, App
598	5	2.7	81	10	US-09-867-550-540	Sequence 540, App	671	5	2.7	114	10	US-09-864-761-46649	Sequence 46649, A
599	5	2.7	82	9	US-10-091-504-678	Sequence 678, App	672	5	2.7	114	10	US-09-867-550-14	Sequence 14, Appl
600	5	2.7	82	10	US-09-764-869-678	Sequence 678, App	673	5	2.7	115	9	US-09-798-889-56	Sequence 56, Appl
601	5	2.7	83	10	US-09-764-864-1504	Sequence 1504, Ap	674	5	2.7	116	10	US-09-939-980-304	Sequence 304, Appl
602	5	2.7	84	9	US-10-013-379-18	Sequence 18, Appl	675	5	2.7	117	10	US-09-867-550-336	Sequence 326, App
603	5	2.7	84	9	US-09-925-299-1207	Sequence 1207, Ap	676	5	2.7	118	10	US-09-734-017A-26	Sequence 26, Appl

677	5	2.7	120	12	US-10-096-703-3	Sequence 3, App1	750	5	2.7	133	12	US-10-153-740-11	Sequence 11, App1
678	5	2.7	122	10	US-10-096-703-4	Sequence 4, App1	751	5	2.7	134	9	US-10-116-273-27	Sequence 27, App1
679	5	2.7	120	12	US-09-731-872-431	Sequence 431, App	752	5	2.7	134	9	US-10-091-483-157	Sequence 157, App
680	5	2.7	123	9	US-09-984-245-169	Sequence 169, App	753	5	2.7	134	10	US-09-164-615-5	Sequence 5, App1
681	5	2.7	123	9	US-09-966-262-169	Sequence 169, App	754	5	2.7	134	10	US-09-764-846-157	Sequence 157, App
682	5	2.7	123	9	US-09-983-966-169	Sequence 169, App	755	5	2.7	135	9	US-10-116-273-26	Sequence 26, App1
683	5	2.7	124	9	US-09-992-598-146	Sequence 146, App	756	5	2.7	136	9	US-10-116-273-23	Sequence 23, App1
684	5	2.7	124	9	US-10-060-845-9	Sequence 9, App1	757	5	2.7	136	9	US-09-738-626-6099	Sequence 6099, App
685	5	2.7	124	9	US-09-989-293A-146	Sequence 146, App	758	5	2.7	136	9	US-10-101-464A-675	Sequence 675, App
686	5	2.7	124	9	US-09-989-735-146	Sequence 146, App	759	5	2.7	136	10	US-09-789-561-101	Sequence 101, App
687	5	2.7	124	9	US-09-990-444-146	Sequence 146, App	760	5	2.7	137	10	US-09-955-286-2	Sequence 2, App1
688	5	2.7	124	9	US-09-988-730-146	Sequence 146, App	761	5	2.7	137	9	US-10-116-273-24	Sequence 24, App1
689	5	2.7	124	9	US-09-990-436-146	Sequence 146, App	762	5	2.7	138	9	US-09-764-868-683	Sequence 683, App
690	5	2.7	124	9	US-09-991-181-146	Sequence 146, App	763	5	2.7	138	9	US-10-116-273-23	Sequence 23, App1
691	5	2.7	124	9	US-09-993-687-146	Sequence 146, App	764	5	2.7	138	10	US-09-912-020-330	Sequence 300, App
692	5	2.7	124	9	US-09-989-734-146	Sequence 146, App	765	5	2.7	139	9	US-10-116-273-22	Sequence 22, App1
693	5	2.7	124	9	US-09-997-653-146	Sequence 146, App	766	5	2.7	140	9	US-10-116-273-21	Sequence 21, App1
694	5	2.7	124	9	US-09-993-667-146	Sequence 146, App	767	5	2.7	140	9	US-10-116-273-34	Sequence 34, App1
695	5	2.7	124	9	US-09-990-438-146	Sequence 146, App	768	5	2.7	140	10	US-09-864-761-41841	Sequence 41841, A
696	5	2.7	124	9	US-09-990-562-146	Sequence 146, App	769	5	2.7	140	10	US-09-731-872-470	Sequence 470, App
697	5	2.7	124	9	US-09-997-428-146	Sequence 146, App	770	5	2.7	141	9	US-10-116-273-20	Sequence 20, App1
698	5	2.7	124	9	US-09-997-666-146	Sequence 146, App	771	5	2.7	141	9	US-10-101-464A-91	Sequence 91, App1
699	5	2.7	124	9	US-09-990-711-146	Sequence 146, App	772	5	2.7	141	10	US-09-730-857-29	Sequence 29, App1
700	5	2.7	124	9	US-09-989-726-146	Sequence 146, App	773	5	2.7	141	10	US-09-730-857-41	Sequence 41, App1
701	5	2.7	124	9	US-09-990-437-146	Sequence 146, App	774	5	2.7	141	10	US-09-730-857-45	Sequence 45, App1
702	5	2.7	124	9	US-09-998-156-146	Sequence 146, App	775	5	2.7	141	10	US-09-730-857-49	Sequence 49, App1
703	5	2.7	124	9	US-09-991-157-146	Sequence 146, App	776	5	2.7	141	10	US-09-730-857-51	Sequence 51, App1
704	5	2.7	124	9	US-09-991-172-146	Sequence 146, App	777	5	2.7	141	10	US-09-730-857-55	Sequence 55, App1
705	5	2.7	124	9	US-09-997-514-146	Sequence 146, App	778	5	2.7	141	10	US-09-730-857-59	Sequence 59, App1
706	5	2.7	124	9	US-09-997-513-146	Sequence 146, App	779	5	2.7	141	10	US-09-730-857-63	Sequence 63, App1
707	5	2.7	124	9	US-09-990-443-146	Sequence 146, App	780	5	2.7	141	10	US-09-730-857-65	Sequence 65, App1
708	5	2.7	124	9	US-09-990-726-146	Sequence 146, App	781	5	2.7	142	9	US-10-116-273-19	Sequence 19, App1
709	5	2.7	124	9	US-09-997-559-146	Sequence 146, App	782	5	2.7	143	9	US-10-116-273-17	Sequence 17, App1
710	5	2.7	124	9	US-09-997-601-146	Sequence 146, App	783	5	2.7	143	9	US-09-738-626-3581	Sequence 3581, App
711	5	2.7	124	9	US-09-989-729A-146	Sequence 146, App	784	5	2.7	143	10	US-09-764-877-1657	Sequence 1657, App
712	5	2.7	124	9	US-09-990-440-146	Sequence 146, App	785	5	2.7	146	10	US-09-815-242-15340	Sequence 5340, App
713	5	2.7	124	9	US-09-991-854-146	Sequence 146, App	786	5	2.7	146	10	US-09-815-242-12225	Sequence 12225, A
714	5	2.7	124	9	US-09-997-349-146	Sequence 146, App	787	5	2.7	146	10	US-09-815-242-11826	Sequence 11826, A
715	5	2.7	124	9	US-09-997-440-146	Sequence 146, App	788	5	2.7	146	10	US-09-815-242-11049	Sequence 11049, A
716	5	2.7	124	9	US-09-997-628-146	Sequence 146, App	789	5	2.7	147	10	US-09-864-761-48692	Sequence 48692, A
717	5	2.7	124	9	US-09-997-683-146	Sequence 146, App	790	5	2.7	147	10	US-09-925-301-1129	Sequence 1129, App
718	5	2.7	124	10	US-09-989-722-146	Sequence 146, App	791	5	2.7	147	10	US-09-939-980-498	Sequence 498, App
719	5	2.7	124	10	US-09-989-723-146	Sequence 146, App	792	5	2.7	147	10	US-09-965-703-44	Sequence 44, App1
720	5	2.7	124	10	US-09-989-729-146	Sequence 146, App	793	5	2.7	148	1	US-08-781-986A-5210	Sequence 5210, App
721	5	2.7	124	10	US-09-989-727-146	Sequence 146, App	794	5	2.7	148	10	US-09-815-242-5673	Sequence 5673, App
722	5	2.7	124	10	US-09-989-731-146	Sequence 146, App	795	5	2.7	148	10	US-09-815-242-12681	Sequence 12681, App
723	5	2.7	124	10	US-09-989-732-146	Sequence 146, App	796	5	2.7	148	10	US-09-815-242-13001	Sequence 13001, A
724	5	2.7	124	10	US-09-991-073-146	Sequence 146, App	797	5	2.7	149	9	US-09-902-941-1901	Sequence 1901, App
725	5	2.7	124	10	US-09-990-442-146	Sequence 146, App	798	5	2.7	149	9	US-09-738-626-6936	Sequence 6936, App
726	5	2.7	124	10	US-09-991-163-146	Sequence 146, App	799	5	2.7	149	9	US-09-849-626-1901	Sequence 1901, App
727	5	2.7	124	10	US-09-993-604-146	Sequence 146, App	800	5	2.7	149	9	US-10-091-483-200	Sequence 200, App
728	5	2.7	124	10	US-09-990-456-146	Sequence 146, App	801	5	2.7	149	9	US-10-017-754-1901	Sequence 1901, App
729	5	2.7	124	10	US-09-989-721-146	Sequence 146, App	802	5	2.7	149	10	US-09-864-761-44466	Sequence 40466, A
730	5	2.7	124	10	US-09-764-877-1235	Sequence 1235, App	803	5	2.7	149	10	US-09-764-846-200	Sequence 200, App
731	5	2.7	124	12	US-10-072-159-9	Sequence 9, App1	804	5	2.7	151	9	US-09-991-496-20	Sequence 20, App1
732	5	2.7	125	9	US-10-011-585A-220	Sequence 220, App	805	5	2.7	151	10	US-09-874-923-20	Sequence 20, App1
733	5	2.7	126	9	US-09-738-626-5169	Sequence 5169, App	806	5	2.7	153	9	US-10-108-605-85	Sequence 85, App1
734	5	2.7	127	10	US-09-922-138-13	Sequence 13, App1	807	5	2.7	154	10	US-09-864-761-34493	Sequence 34493, A
735	5	2.7	128	9	US-10-116-273-33	Sequence 33, App1	808	5	2.7	155	9	US-10-101-464A-487	Sequence 497, App
736	5	2.7	128	10	US-09-815-242-10887	Sequence 10887, A	809	5	2.7	156	12	US-10-027-450-6	Sequence 6, App1
737	5	2.7	129	9	US-10-116-273-32	Sequence 32, App1	810	5	2.7	157	9	US-10-131-406-2	Sequence 2, App1
738	5	2.7	130	9	US-10-116-273-31	Sequence 31, App1	811	5	2.7	157	10	US-09-864-761-33245	Sequence 33245, A
739	5	2.7	130	9	US-09-983-802-633	Sequence 633, App	812	5	2.7	158	9	US-10-076-622-508	Sequence 508, App
740	5	2.7	130	9	US-10-101-464A-496	Sequence 496, App	813	5	2.7	158	10	US-09-734-569-76	Sequence 569, App1
741	5	2.7	130	10	US-09-815-242-4999	Sequence 4999, App	814	5	2.7	158	12	US-10-007-806-558	Sequence 508, App
742	5	2.7	130	10	US-09-815-242-5052	Sequence 5052, App	815	5	2.7	159	9	US-09-738-626-5261	Sequence 5261, App
743	5	2.7	131	9	US-10-116-273-30	Sequence 30, App1	816	5	2.7	160	10	US-09-771-209-28	Sequence 28, App1
744	5	2.7	131	10	US-09-925-301-1552	Sequence 1552, App	817	5	2.7	161	9	US-10-101-464A-676	Sequence 676, App1
745	5	2.7	132	9	US-10-116-273-29	Sequence 29, App1	818	5	2.7	161	10	US-09-922-261-443	Sequence 423, App
746	5	2.7	132	10	US-09-867-550-60	Sequence 60, App1	819	5	2.7	162	9	US-09-798-889-107	Sequence 107, App
747	5	2.7	133	9	US-10-116-273-28	Sequence 28, App1	820	5	2.7	162	9	US-09-975-719-431	Sequence 431, App
748	5	2.7	133	10	US-09-916-790-26	Sequence 26, App1	821	5	2.7	162	9	US-10-101-464A-672	Sequence 672, App
749	5	2.7	133	10	US-09-901-436A-7	Sequence 7, App1	822	5	2.7	163	9	US-09-975-719-335	Sequence 335, App

823	5	2.7	153	10	US-09-815-242-12218	Sequence 12218, A	896	5	2.7	191	9	US-10-101-464A-735	Sequence 735, App
824	5	2.7	154	9	US-09-738-626-4742	Sequence 4742, App	897	5	2.7	191	9	US-10-086-156-98	Sequence 98, Appl
825	5	2.7	164	10	US-09-867-550-614	Sequence 614, Ap	898	5	2.7	191	10	US-09-815-242-5067	Sequence 5067, Ap
826	5	2.7	164	10	US-09-925-300-1298	Sequence 1298, Ap	899	5	2.7	191	10	US-09-947-1316-5	Sequence 5, Appl
827	5	2.7	165	9	US-09-764-868-950	Sequence 950, App	900	5	2.7	193	10	US-09-828-664-91	Sequence 91, Appl
828	5	2.7	165	9	US-09-738-626-5314	Sequence 5314, Ap	901	5	2.7	193	10	US-09-761-288-77	Sequence 77, Appl
829	5	2.7	166	9	US-09-764-868-946	Sequence 946, App	902	5	2.7	193	10	US-09-761-288-83	Sequence 83, Appl
830	5	2.7	166	9	US-10-116-273-18	Sequence 18, Appl	903	5	2.7	194	10	US-09-784-249-4	Sequence 4, Appl
831	5	2.7	166	9	US-09-738-626-6282	Sequence 6282, Ap	904	5	2.7	194	10	US-09-789-561-174	Sequence 174, App
832	5	2.7	166	12	US-10-142-373-4	Sequence 4, Appl	905	5	2.7	194	10	US-09-910-150-28	Sequence 28, Appl
833	5	2.7	167	9	US-10-012-896-481	Sequence 481, App	906	5	2.7	195	9	US-10-125-540-318	Sequence 318, App
834	5	2.7	167	9	US-09-895-793-481	Sequence 481, App	907	5	2.7	195	10	US-09-764-870-318	Sequence 318, App
835	5	2.7	167	9	US-09-895-814-481	Sequence 481, App	908	5	2.7	195	10	US-09-764-885-487	Sequence 487, App
836	5	2.7	167	10	US-09-759-143-481	Sequence 481, App	909	5	2.7	196	10	US-09-733-507-16	Sequence 16, Appl
837	5	2.7	167	10	US-09-780-669-481	Sequence 481, App	910	5	2.7	196	10	US-09-785-738A-6	Sequence 6, Appl
838	5	2.7	167	10	US-09-822-827-481	Sequence 481, App	911	5	2.7	196	10	US-09-867-550-1892	Sequence 1892, Ap
839	5	2.7	168	9	US-09-764-868-957	Sequence 957, App	912	5	2.7	196	9	US-09-738-626-4812	Sequence 4812, Ap
840	5	2.7	168	9	US-09-764-868-1019	Sequence 1019, Ap	913	5	2.7	198	10	US-09-785-738A-4	Sequence 4, Appl
841	5	2.7	168	9	US-10-025-380-198	Sequence 198, App	914	5	2.7	199	9	US-09-866-050A-230	Sequence 230, App
842	5	2.7	168	9	US-10-232-188-3	Sequence 3, Appl	915	5	2.7	199	8	US-08-424-550B-395	Sequence 395, App
843	5	2.7	168	9	US-09-820-843A-49	Sequence 49, Appl	916	5	2.7	200	10	US-09-785-738A-12	Sequence 12, Appl
844	5	2.7	168	10	US-09-791-118A-6	Sequence 6, Appl	917	5	2.7	200	10	US-09-928-412-2	Sequence 2, Appl
845	5	2.7	168	10	US-09-922-217-198	Sequence 198, App	918	5	2.7	201	9	US-10-163-499-2	Sequence 2, Appl
846	5	2.7	168	10	US-09-833-263-198	Sequence 198, App	919	5	2.7	201	9	US-10-163-499-4	Sequence 4, Appl
847	5	2.7	169	9	US-10-136-511-2	Sequence 2, Appl	920	5	2.7	201	9	US-09-815-242-5282	Sequence 5282, Ap
848	5	2.7	169	10	US-09-345-790-2	Sequence 2, Appl	921	5	2.7	202	9	US-09-847-102A-65	Sequence 65, Appl
849	5	2.7	170	9	US-10-101-464A-527	Sequence 527, App	922	5	2.7	202	9	US-10-101-464A-537	Sequence 537, App
850	5	2.7	170	10	US-09-734-569-102	Sequence 102, App	923	5	2.7	203	10	US-09-823-901-7	Sequence 7, Appl
851	5	2.7	170	10	US-09-881-752A-348	Sequence 348, App	924	5	2.7	205	9	US-09-989-920-228	Sequence 228, App
852	5	2.7	172	9	US-10-087-195-6	Sequence 6, Appl	925	5	2.7	205	9	US-09-927-827-63	Sequence 63, Appl
853	5	2.7	172	9	US-09-738-626-6799	Sequence 6799, Ap	926	5	2.7	205	10	US-09-921-397-87	Sequence 87, Appl
854	5	2.7	172	12	US-10-085-572-6	Sequence 6, Appl	927	5	2.7	206	9	US-09-997-816-4	Sequence 4, Appl
855	5	2.7	173	10	US-09-205-658-216	Sequence 216, Appl	928	5	2.7	206	10	US-09-864-761-33643	Sequence 33643, A
856	5	2.7	174	10	US-09-861-451A-58	Sequence 58, Appl	929	5	2.7	206	10	US-09-828-310-13	Sequence 13, Appl
857	5	2.7	176	9	US-10-036-729-2	Sequence 2, Appl	930	5	2.7	206	10	US-09-800-722-149	Sequence 149, App
858	5	2.7	176	9	US-10-013-379-5	Sequence 5, Appl	931	5	2.7	207	10	US-09-925-301-1160	Sequence 1160, Ap
859	5	2.7	177	9	US-10-151-882-39	Sequence 39, Appl	932	5	2.7	207	10	US-09-800-729-113	Sequence 113, App
860	5	2.7	177	10	US-09-272-809-15	Sequence 15, Appl	933	5	2.7	207	10	US-09-965-602-34	Sequence 34, Appl
861	5	2.7	177	12	US-10-080-455-1	Sequence 1, Appl	934	5	2.7	207	10	US-09-995-598-28	Sequence 28, Appl
862	5	2.7	177	12	US-10-116-378-16	Sequence 16, Appl	935	5	2.7	208	9	US-09-870-759-62	Sequence 62, Appl
863	5	2.7	178	9	US-09-764-868-792	Sequence 792, App	936	5	2.7	209	9	US-09-870-759-63	Sequence 63, Appl
864	5	2.7	178	9	US-09-738-626-5511	Sequence 5511, Ap	937	5	2.7	210	10	US-09-761-534A-8	Sequence 8, Appl
865	5	2.7	178	9	US-09-738-626-5841	Sequence 5841, Ap	938	5	2.7	211	9	US-10-222-577-5	Sequence 5, Appl
866	5	2.7	178	10	US-09-780-717-20	Sequence 20, Appl	939	5	2.7	211	9	US-10-222-578-5	Sequence 5, Appl
867	5	2.7	178	10	US-09-815-242-10448	Sequence 10448, A	940	5	2.7	211	10	US-09-790-045-5	Sequence 5, Appl
868	5	2.7	179	9	US-09-999-686-37	Sequence 37, Appl	941	5	2.7	212	9	US-09-975-719-235	Sequence 235, App
869	5	2.7	179	10	US-09-815-242-5534	Sequence 5534, Ap	942	5	2.7	212	10	US-09-749-728B-65	Sequence 65, Appl
870	5	2.7	180	10	US-09-910-150-29	Sequence 29, Appl	943	5	2.7	213	9	US-09-984-245-237	Sequence 237, App
871	5	2.7	180	10	US-09-771-730-131	Sequence 131, App	944	5	2.7	213	9	US-09-966-262-237	Sequence 237, App
872	5	2.7	181	9	US-09-738-626-5074	Sequence 5074, Ap	945	5	2.7	213	9	US-09-983-966-237	Sequence 237, App
873	5	2.7	181	10	US-09-780-717-23	Sequence 23, Appl	946	5	2.7	213	10	US-09-844-468-4	Sequence 4, Appl
874	5	2.7	181	10	US-09-206-639-1	Sequence 1, Appl	947	5	2.7	213	10	US-09-811-288-229	Sequence 239, App
875	5	2.7	182	9	US-10-080-960-26	Sequence 26, Appl	948	5	2.7	214	9	US-10-005-057A-35	Sequence 35, Appl
876	5	2.7	182	9	US-09-975-719-327	Sequence 327, App	949	5	2.7	214	9	US-10-125-452-22	Sequence 22, Appl
877	5	2.7	182	9	US-10-163-499-6	Sequence 6, Appl	950	5	2.7	214	9	US-09-955-504-22	Sequence 22, Appl
878	5	2.7	182	9	US-10-163-499-13	Sequence 13, Appl	951	5	2.7	214	10	US-09-924-256A-0	Sequence 30, Appl
879	5	2.7	182	10	US-09-815-242-5588	Sequence 5588, Ap	952	5	2.7	215	10	US-09-825-414-60	Sequence 60, Appl
880	5	2.7	183	9	US-10-116-255-27	Sequence 27, Appl	953	5	2.7	216	10	US-09-864-761-38168	Sequence 38168, A
881	5	2.7	183	10	US-09-815-242-5833	Sequence 5833, Ap	954	5	2.7	216	10	US-09-771-730-50	Sequence 50, Appl
882	5	2.7	186	10	US-09-921-397-88	Sequence 88, Appl	955	5	2.7	216	10	US-09-771-730-143	Sequence 143, App
883	5	2.7	187	9	US-10-101-482-17	Sequence 17, Appl	956	5	2.7	216	10	US-09-747-155-427	Sequence 146, App
884	5	2.7	187	10	US-09-986-480-368	Sequence 368, App	957	5	2.7	216	10	US-09-747-155-427	Sequence 427, App
885	5	2.7	187	10	US-09-864-761-35716	Sequence 35716, A	958	5	2.7	217	9	US-10-108-605-273	Sequence 273, App
886	5	2.7	187	10	US-09-864-761-36096	Sequence 36096, A	959	5	2.7	217	9	US-09-738-626-4573	Sequence 4573, Ap
887	5	2.7	188	9	US-09-791-932-101	Sequence 101, App	960	5	2.7	217	10	US-09-925-301-1048	Sequence 1048, Ap
888	5	2.7	188	10	US-09-193-538-17	Sequence 17, Appl	961	5	2.7	217	10	US-09-815-242-11540	Sequence 11540, A
889	5	2.7	188	10	US-09-815-242-12426	Sequence 12426, A	962	5	2.7	218	10	US-09-925-300-1571	Sequence 1571, Ap
890	5	2.7	188	10	US-09-815-242-12773	Sequence 12773, A	963	5	2.7	219	10	US-09-815-242-10140	Sequence 10140, A
891	5	2.7	188	10	US-09-815-242-13139	Sequence 13139, A	964	5	2.7	220	10	US-09-815-242-10428	Sequence 10428, A
892	5	2.7	190	10	US-09-815-242-10092	Sequence 10092, A	965	5	2.7	220	10	US-09-815-242-11614	Sequence 11614, A
893	5	2.7	191	9	US-09-860-670-110	Sequence 110, App	966	5	2.7	220	10	US-09-923-304-2	Sequence 2, Appl
894	5	2.7	191	9	US-10-040-805-18	Sequence 18, Appl	967	5	2.7	221	10	US-09-995-558B-14	Sequence 14, Appl
895	5	2.7	191	9	US-10-080-980-64	Sequence 64, Appl	968	5	2.7	222	10	US-09-925-300-1639	Sequence 1639, Ap



```
969 5 2.7 223 9 US-09-738-626-3860 Sequence 3860, Ap
970 5 2.7 223 10 US-09-646-5908-16 Sequence 16, Appl
971 5 2.7 223 10 US-09-993-170-3 Sequence 3, Appl
972 5 2.7 224 10 US-09-993-170-16 Sequence 16, Appl
973 5 2.7 225 10 US-09-922-261-415 Sequence 415, Ap
974 5 2.7 226 9 US-09-738-626-3665 Sequence 3665, Ap
975 5 2.7 226 10 US-09-768-826-52 Sequence 52, Appl
976 5 2.7 228 9 US-09-866-0504-684 Sequence 684, App
977 5 2.7 229 9 US-10-079-623-359 Sequence 359, App
978 5 2.7 229 9 US-10-043-487-252 Sequence 252, App
979 5 2.7 230 9 US-09-738-626-5076 Sequence 5076, Ap
980 5 2.7 230 9 US-10-006-922-8 Sequence 8, Appl
981 5 2.7 230 9 US-10-006-922-30 Sequence 30, Appl
982 5 2.7 230 9 US-10-006-922-34 Sequence 34, Appl
983 5 2.7 230 9 US-10-081-864-6 Sequence 6, Appl
984 5 2.7 230 9 US-10-081-864-16 Sequence 16, Appl
985 5 2.7 230 9 US-10-081-864-18 Sequence 18, Appl
986 5 2.7 230 9 US-09-776-724A-147 Sequence 147, App
987 5 2.7 230 10 US-09-995-598-38 Sequence 38, Appl
988 5 2.7 231 9 US-10-006-922-31 Sequence 31, Appl
989 5 2.7 231 9 US-10-006-922-32 Sequence 32, Appl
990 5 2.7 232 9 US-09-738-626-4476 Sequence 4476, Ap
991 5 2.7 232 9 US-09-975-719-37 Sequence 37, Appl
992 5 2.7 232 10 US-09-794-975-7 Sequence 7, Appl
993 5 2.7 232 10 US-09-825-561A-4 Sequence 4, Appl
994 5 2.7 232 10 US-09-925-300-1695 Sequence 1695, Ap
995 5 2.7 234 10 US-09-924-256A-52 Sequence 52, Appl
996 5 2.7 235 9 US-09-792-251-2 Sequence 2, Appl
997 5 2.7 235 9 US-10-076-622-487 Sequence 487, App
998 5 2.7 235 9 US-10-081-051-66 Sequence 66, Appl
999 5 2.7 235 10 US-09-925-302-542 Sequence 542, App
1000 5 2.7 235 10 US-09-815-242-13256 Sequence 13256, A
```

## ALIGNMENTS

```
RESULT 1
US-09-987-190-6
; Sequence 6, Application US/09987190
; Patent No. US20020058293A1
; GENERAL INFORMATION:
; APPLICANT: TAKESAKO, Kazutosh
; APPLICANT: MIZUTANI, Shigetoshi
; APPLICANT: ENDO, Masahito
; APPLICANT: KATO, Ikunoshi
; TITLE OF INVENTION: FUNGAL ANTIGENS AND PROCESS FOR PRODUCING THE SAME
; FILE REFERENCE: 1422-0502P
; CURRENT APPLICATION NUMBER: US/09/987,190
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: 09/262,856
; PRIOR FILING DATE: 1999-03-04
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-987-190-6

Query Match 100.0%; Score 188; DB 10; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.8e-173;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ATEPITGQNMKEIHFKHQTYYNNLNLSIRQAVAKSGEVKKLVALQKAINFGGGYL 60
DB 1 ATEPITGQNMKEIHFKHQTYYNNLNLSIRQAVAKSGEVKKLVALQKAINFGGGYL 60
QY 61 NHCLMWMKLVAVSHSGGQPSSEDSKLGKQIVKQFSLDKLLEITNGKLAGIGSGMARIIV 120
DB 61 NHCLMWMKLVAVSHSGGQPSSEDSKLGKQIVKQFSLDKLLEITNGKLAGIGSGMARIIV 120
QY 121 NKKANGTIDIVITTAADVTDLNLVPLIAIDAWKHAIVLYOYONVKADYFKNLMHVINMK 180
```

```
DB 121 NKKANGTIDIVITTAADVTDLNLVPLIAIDAWKHAIVLYOYONVKADYFKNLMHVINMK 180
QY 181 EAERREF 188
DB 181 EAERREF 188
```

```
RESULT 2
US-09-727-855B-5
; Sequence 5, Application US/09727855B
; Patent No. US20020168703A1
; GENERAL INFORMATION:
; APPLICANT: HOSHINO, Tatsuo
; APPLICANT: OJIMA, Kazuyuki
; APPLICANT: SETOGUCHI, Yutaka
; TITLE OF INVENTION: PROCESS FOR THE MANUFACTURE OF CAROTENOIDS AND BIOLOGICALLY USEF
; FILE REFERENCE: C38435/111694
; CURRENT APPLICATION NUMBER: US/09/727,855B
; CURRENT FILING DATE: 2000-12-01
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Phaffia rhodozyma
US-09-727-855B-5
```

```
Query Match 4.8%; Score 9; DB 9; Length 222;
Best Local Similarity 100.0%; Pred. No. 0.43;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 16 TKXHQTYVN 24
DB 53 TKXHQTYVN 61
```

```
RESULT 3
US-10-101-464A-721
; Sequence 721, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 721
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Pinus radiata
US-10-101-464A-721
```

```
Query Match 3.7%; Score 7; DB 9; Length 121;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 95 GSDLKL 101
```



Db 98 GSLDKL1 104

```
RESULT 4
US-09-727-855B-7
; Sequence 7, Application US/09727855B
; Patent No. US20020168703A1
; GENERAL INFORMATION:
; APPLICANT: HOSHINO, Tatsuo
; APPLICANT: OJIMA, Kazuyuki
; APPLICANT: SETOGUCHI, Yutaka
; TITLE OF INVENTION: PROCESS FOR THE MANUFACTURE OF CAROTENOIDS AND BIOLOGICALLY USEFUL
; FILE REFERENCE: C38435/111694
; CURRENT FILING DATE: 2000-12-01
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Phaffia rhodozyma
US-09-727-855B-7
```

```
Query Match
Best Local Similarity 100.0%; Score 7; DB 9; Length 198;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 17 KKHOTYV 23
Db 31 KKHOTYV 37
```

```
RESULT 5
US-10-101-464A-702
; Sequence 702, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; FILE REFERENCE: 11000.1020c2
; CURRENT FILING DATE: 2002-03-18
; PRIOR FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: US/10/101,464A
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 702
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Pinus radiata
US-10-101-464A-702
```

```
Query Match
Best Local Similarity 100.0%; Score 7; DB 9; Length 238;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 95 GSLDKL1 101
Db 140 GSLDKL1 146
```

```
RESULT 6
US-09-712-363-271
```

```
; Sequence 271, Application US/09712363
; Patent No. US20020164588A1
; GENERAL INFORMATION:
; APPLICANT: Eisenberg, David
; APPLICANT: Rotstein, Sergio H.
; APPLICANT: Marcotte, Edward M.
; TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
; FILE REFERENCE: 07419-032001
; CURRENT FILING DATE: US/09/712,363
; CURRENT FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: PCT/US00/02246
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,531
; PRIOR FILING DATE: 2000-02-01
; PRIOR APPLICATION NUMBER: 60/117,844
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: 60/118,206,
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: 60/126,593
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 60/134,093
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/134,092
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/165,124
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/165,086
; PRIOR FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 292
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 271
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-712-363-271
```

```
Query Match
Best Local Similarity 100.0%; Score 7; DB 9; Length 329;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 77 GQPPSED 83
Db 207 GQPPSED 213
```

```
RESULT 7
US-09-815-242-11129
; Sequence 1129, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Onlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT FILING DATE: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
```

```
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11129
; LENGTH: 506
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-815-242-11129
```

```
Query Match      3.7%; Score 7; DB 10; Length 506;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 105 NGKLAGI 111
Db 481 NGKLAGI 487
```

```
RESULT 8
US-09-817-913-13
; Sequence 13, Application US/09817913
; Patent No. US20020061860A1
; GENERAL INFORMATION:
; APPLICANT: Li, Zuomei
; APPLICANT: Bonfils, Claire
; APPLICANT: Besteman, Jeffrey
; TITLE OF INVENTION: Inhibition of Specific Histone Deacetylase Isoforms
; FILE REFERENCE: 106101.145
; CURRENT APPLICATION NUMBER: US/09/817,913
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/192,157
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 855
; TYPE: PRT
; ORGANISM: Human
US-09-817-913-13
```

```
Query Match      3.7%; Score 7; DB 10; Length 855;
Best Local Similarity 100.0%; Pred. No. 11e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 105 NGKLAGI 111
Db 561 NGKLAGI 567
```

```
RESULT 9
US-09-817-538-13
; Sequence 13, Application US/09817538
; Patent No. US2002013762A1
; GENERAL INFORMATION:
; APPLICANT: Li, Zuomei
; APPLICANT: Bonfils, Claire
; APPLICANT: Besteman, Jeffrey
; TITLE OF INVENTION: Antisense Oligonucleotide Inhibition of Specific Histone
; FILE REFERENCE: 106101.144
; CURRENT APPLICATION NUMBER: US/09/817,538
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/192,157
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 855
; TYPE: PRT
```

```
; ORGANISM: Human
US-09-817-538-13
```

```
Query Match      3.7%; Score 7; DB 10; Length 855;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 105 NGKLAGI 111
Db 561 NGKLAGI 567
```

```
RESULT 10
US-09-738-626-4048
; Sequence 4048, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIRO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4048
; LENGTH: 1333
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4048
```

```
Query Match      3.7%; Score 7; DB 9; Length 1333;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 38 SKGEVKK 44
Db 24 SKGEVKK 30
```

```
RESULT 11
US-09-853-830-28
; Sequence 28, Application US/09853830
; Patent No. US20020107388A1
; GENERAL INFORMATION:
; APPLICANT: Vandendark, Arthur A.
; TITLE OF INVENTION: Methods of Identifying and Monitoring
; FILE REFERENCE: P-IM 4734
; CURRENT APPLICATION NUMBER: US/09/853,830
; PRIOR FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 184
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-853-830-28
```

Query Match 3.2%; Score 6; DB 10; Length 21;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 40 GYVKL 45  
16 GYVKL 21  
DB

RESULT 12  
US-09-987-190-2  
Sequence 2, Application US/09987190  
Patent No. US20020058293A1  
GENERAL INFORMATION:  
APPLICANT: TAKESAKO, Kazutoh  
APPLICANT: MIZUTANI, Shigetoshi  
APPLICANT: ENDO, Masahiro  
APPLICANT: KATO, Ikunoshin  
TITLE OF INVENTION: FUNGAL ANTIGENS AND PROCESS FOR PRODUCING THE SAME  
FILE REFERENCE: 1422-0502P  
CURRENT APPLICATION NUMBER: US/09/987,190  
CURRENT FILING DATE: 2001-11-13  
PRIOR APPLICATION NUMBER: 09/262,856  
PRIOR FILING DATE: 1999-03-04  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 2  
LENGTH: 30  
TYPE: PRT  
ORGANISM: Candida albicans  
FEATURE:  
NAME/KEY: Unsure  
LOCATION: (1)..(30)  
OTHER INFORMATION: any Xaa = any amino acid, unknown or other  
US-09-987-190-2

Query Match 3.2%; Score 6; DB 10; Length 30;  
Best Local Similarity 100.0%; Pred. No. 60;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATEPYI 6  
13 ATEPYI 18  
DB

RESULT 13  
US-09-864-761-36462  
Sequence 36462, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aecomica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 36462  
LENGTH: 39  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC011257.2  
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 0.95  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.3  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.97  
OTHER INFORMATION: EST HUMAN HIT: BR219182.1, EVALUE 3.20e+00  
NAME/KEY: unsure  
LOCATION: 24  
US-09-864-761-36462

Query Match 3.2%; Score 6; DB 10; Length 39;  
Best Local Similarity 100.0%; Pred. No. 75;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 GKQIVK 92  
16 GKQIVK 21  
DB

RESULT 14  
US-09-981-876-166  
Sequence 166, Application US/09981876  
Patent No. US20020164669A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 70 Human Secreted Proteins  
FILE REFERENCE: P2001P1  
CURRENT APPLICATION NUMBER: US/09/981,876  
CURRENT FILING DATE: 2001-10-19  
PRIOR APPLICATION NUMBER: 09/148,545  
PRIOR FILING DATE: 1998-09-04  
PRIOR APPLICATION NUMBER: 60/040,162  
PRIOR FILING DATE: 1997-03-07  
PRIOR APPLICATION NUMBER: 60/040,333  
PRIOR FILING DATE: 1997-03-07  
PRIOR APPLICATION NUMBER: 60/038,621  
PRIOR FILING DATE: 1997-03-07  
PRIOR APPLICATION NUMBER: 60/040,161

Page 12

[illegible]

```

; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/043,578
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,576
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/047,501
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/043,670
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/056,632
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,664
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,876
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,881
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,909
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,875
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,862
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,887
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,908
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/048,964
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/057,650
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/056,884
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 280
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 166
; LENGTH: 81
```

```

Query Match      3.2%; Score 6; DB 9; Length 81;
Best Local Similarity 100.0%; Pred.No.1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 ATEPYI 6
      |||||
Db      71 ATEPYI 76
```

```

SULT 15
-09-148-545-166
; Sequence 166, Application US/09148545
; Publication No. US20030027132A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: P2001PI
; CURRENT APPLICATION NUMBER: US/09/148,545
; CURRENT FILING DATE: 1998-09-04
; EARLIER APPLICATION NUMBER: PCT/US98/04482
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,161
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
```

```

; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
```

EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 877  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 889  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 893  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 630  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 878  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 662  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 872  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 882  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 637  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 903  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 886  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 879  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 880  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 894  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 911  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 636  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 874  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 910  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 864  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 631  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 845  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 892  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/047, 595  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/057, 761  
EARLIER FILING DATE: 05-Sep-1997  
EARLIER APPLICATION NUMBER: 60/047, 599  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047, 588  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047, 585  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047, 586  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047, 590  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047, 594  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047, 589  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047, 593  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047, 614  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043, 578  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043, 576  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/047, 501  
EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/043, 670  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/056, 632  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 664  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 876  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 881  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 909  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 875  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 862  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 887  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056, 908  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/048, 964  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/057, 650  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/056, 884  
EARLIER FILING DATE: 1997-08-22  
NUMBER OF SEQ ID NOS: 280  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 166  
LENGTH: 81

Query Match 3.2%; Score 6; DB 9; Length 81;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATEPYI 6  
|||||  
Db 71 ATEPYI 76

RESULT 16  
US-09-867-550-1282  
Sequence 1282, Application US/09867550  
Patent No. US20020082206A1  
GENERAL INFORMATION:  
APPLICANT: Leach, Martin D.  
APPLICANT: Mehraban, Ruad,  
APPLICANT: Conley, Pamela  
APPLICANT: Law, Debbie  
TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells and  
TITLE OF INVENTION: Thereby  
FILE REFERENCE: 21402-013 (Cura-313)  
CURRENT APPLICATION NUMBER: US/09/867,550  
CURRENT FILING DATE: 2001-09-20  
PRIOR APPLICATION NUMBER: USSN 60/208,427  
PRIOR FILING DATE: 2000-05-30  
NUMBER OF SEQ ID NOS: 2125  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1282  
LENGTH: 84  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (7)  
OTHER INFORMATION: Wherein Xaa may be any one of Ile or Leu or Phe or Val  
NAME/KEY: VARIANT  
LOCATION: (80)  
OTHER INFORMATION: wherein Xaa may be any one of Ile or Leu or Val  
US-09-867-550-1282

Query Match 3.2%; Score 6; DB 10; Length 84;

Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 81 SEDSKL 86  
Db 68 SEDSKL 73

RESULT 17  
US-09-993-333-12  
; Sequence 12, Application US/09993333  
; Patent No. US20020156040A1  
; GENERAL INFORMATION:  
; APPLICANT: Oberley, Larry Wayne  
; APPLICANT: Weydert, Christine J.  
; APPLICANT: Smith, Benjamin Barnes  
; TITLE OF INVENTION: Reduction of antioxidant enzyme levels in tumor cells using antisense  
; FILE REFERENCE: 875,042US1  
; CURRENT APPLICATION NUMBER: US/09/993,333  
; PRIOR FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: US 60/248,328  
; PRIOR FILING DATE: 2000-11-14  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12  
; LENGTH: 95  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-993-333-12

Query Match 3.2%; Score 6; DB 9; Length 95;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 YVNNLN 27  
Db 58 YVNNLN 63

RESULT 18  
US-09-515-806-28  
; Sequence 28, Application US/09515806  
; Patent No. US20020132321A1  
; GENERAL INFORMATION:  
; APPLICANT: COOK, WILLIAM J.  
; APPLICANT: KAPILLER-LIBERMAN, ROSANA  
; TITLE OF INVENTION: 14790, NOVEL PROTEIN KINASE MOLECULE AND USES THEREFOR  
; FILE REFERENCE: 38155-20002.00  
; CURRENT APPLICATION NUMBER: US/09/515,806  
; CURRENT FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 28  
; LENGTH: 104  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Consensus  
; OTHER INFORMATION: Kinase sequence  
US-09-515-806-28

Query Match 3.2%; Score 6; DB 10; Length 104;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 74 HGGGQP 79  
Db 45 HGGGQP 50

RESULT 19  
US-09-738-626-4908

; Sequence 4908, Application US/09738626  
; Publication No. US20020197605A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAGAWA, SATOSHI  
; APPLICANT: MIZOGUCHI, HIROSHI  
; APPLICANT: ANDO, SEIKO  
; APPLICANT: HAYASHI, MIKIRO  
; APPLICANT: OCHIAI, KEIKO  
; APPLICANT: YOKOI, HARUHIKO  
; APPLICANT: TATEISHI, NAKO  
; APPLICANT: SENOH, AKIHIRO  
; APPLICANT: IKEDA, MASATO  
; APPLICANT: OZAKI, AKIO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-125  
; CURRENT APPLICATION NUMBER: US/09/738,626  
; PRIOR FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: JP 99/377484  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: JP 00/159162  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: JP 00/280988  
; PRIOR FILING DATE: 2000-08-03  
; NUMBER OF SEQ ID NOS: 7059  
; SOFTWARE: PatentIn ver. 3.0  
; SEQ ID NO 4908  
; LENGTH: 123  
; TYPE: PRT  
; ORGANISM: Corynebacterium glutamicum  
US-09-738-626-4908

Query Match 3.2%; Score 6; DB 9; Length 123;  
Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 LVALQK 50  
Db 52 LVALQK 57

RESULT 20  
US-10-116-273-16  
; Sequence 16, Application US/10116273  
; Publication No. US20020192183A1  
; GENERAL INFORMATION:  
; APPLICANT: Anne Dam Jensen  
; TITLE OF INVENTION: Interferon gamma polypeptide variants  
; FILE REFERENCE: 231us410 - INFG variants  
; CURRENT APPLICATION NUMBER: US/10/116,273  
; CURRENT FILING DATE: 2002-04-04  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 16  
; LENGTH: 128  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: C-terminally  
; OTHER INFORMATION: truncated [S99F]NIFNG  
US-10-116-273-16

Query Match 3.2%; Score 6; DB 9; Length 128;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 TVTDLN 144  
Db 99 TVTDLN 104

RESULT 21  
US-10-116-273-15  
; Sequence 15, Application US/10116273

```

; Publication No. US20020192183A1
; GENERAL INFORMATION:
; APPLICANT: Anne Dam Jensen
; TITLE OF INVENTION: Interferon gamma polypeptide variants
; FILE REFERENCE: 231us410 - INFG variants
; CURRENT APPLICATION NUMBER: US/10/116,273
; CURRENT FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: C-terminally
; OTHER INFORMATION: truncated [S99T]huIFNG
US-10-116-273-15

Query Match
Best Local Similarity 3.2%; Score 6; DB 9; Length 129;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 TVTDLN 144
   |||||
Db 99 TVTDLN 104

RESULT 22
US-10-116-273-14
; Sequence 14, Application US/10116273
; Publication No. US20020192183A1
; GENERAL INFORMATION:
; APPLICANT: Anne Dam Jensen
; TITLE OF INVENTION: Interferon gamma polypeptide variants
; FILE REFERENCE: 231us410 - INFG variants
; CURRENT APPLICATION NUMBER: US/10/116,273
; CURRENT FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 130
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: C-terminally
; OTHER INFORMATION: truncated [S99T]huIFNG
US-10-116-273-14

Query Match
Best Local Similarity 3.2%; Score 6; DB 9; Length 130;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 TVTDLN 144
   |||||
Db 99 TVTDLN 104

RESULT 23
US-10-116-273-13
; Sequence 13, Application US/10116273
; Publication No. US20020192183A1
; GENERAL INFORMATION:
; APPLICANT: Anne Dam Jensen
; TITLE OF INVENTION: Interferon gamma polypeptide variants
; FILE REFERENCE: 231us410 - INFG variants
; CURRENT APPLICATION NUMBER: US/10/116,273
; CURRENT FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 131
; TYPE: PRT
; ORGANISM: Artificial Sequence
```

```

; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: C-terminally
; OTHER INFORMATION: truncated [S99T]huIFNG
US-10-116-273-13

Query Match
Best Local Similarity 3.2%; Score 6; DB 9; Length 131;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 TVTDLN 144
   |||||
Db 99 TVTDLN 104

RESULT 24
US-10-116-273-12
; Sequence 12, Application US/10116273
; Publication No. US20020192183A1
; GENERAL INFORMATION:
; APPLICANT: Anne Dam Jensen
; TITLE OF INVENTION: Interferon gamma polypeptide variants
; FILE REFERENCE: 231us410 - INFG variants
; CURRENT APPLICATION NUMBER: US/10/116,273
; CURRENT FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: C-terminally
; OTHER INFORMATION: truncated [S99T]huIFNG
US-10-116-273-12

Query Match
Best Local Similarity 3.2%; Score 6; DB 9; Length 132;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 TVTDLN 144
   |||||
Db 99 TVTDLN 104

RESULT 25
US-10-116-273-11
; Sequence 11, Application US/10116273
; Publication No. US20020192183A1
; GENERAL INFORMATION:
; APPLICANT: Anne Dam Jensen
; TITLE OF INVENTION: Interferon gamma polypeptide variants
; FILE REFERENCE: 231us410 - INFG variants
; CURRENT APPLICATION NUMBER: US/10/116,273
; CURRENT FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 133
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: C-terminally
; OTHER INFORMATION: truncated [S99T]huIFNG
US-10-116-273-11

Query Match
Best Local Similarity 3.2%; Score 6; DB 9; Length 133;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 TVTDLN 144
   |||||
Db 99 TVTDLN 104
```



```
RESULT 26
US-10-116-273-10
; Sequence 10, Application US/10116273
; Publication No. US20020192183A1
; GENERAL INFORMATION:
; APPLICANT: Anne Dam Jensen
; TITLE OF INVENTION: Interferon gamma polypeptide variants
; FILE REFERENCE: 231us410 - INFG variants
; CURRENT APPLICATION NUMBER: US/10/116,273
; CURRENT FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 10
; LENGTH: 134
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: C-terminally
; OTHER INFORMATION: truncated [S99T]huIFNG
; -10-116-273-10

Query Match
3.2%; Score 6; DB 9; Length 134;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 TVTDLN 144
|||||
Db 99 TVTDLN 104

RESULT 27
US-10-116-273-9
; Sequence 9, Application US/10116273
; Publication No. US20020192183A1
; GENERAL INFORMATION:
; APPLICANT: Anne Dam Jensen
; TITLE OF INVENTION: Interferon gamma polypeptide variants
; FILE REFERENCE: 231us410 - INFG variants
; CURRENT APPLICATION NUMBER: US/10/116,273
; CURRENT FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 9
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: C-terminally
; OTHER INFORMATION: truncated [S99T]huIFNG
; -10-116-273-9

Query Match
3.2%; Score 6; DB 9; Length 135;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 TVTDLN 144
|||||
Db 99 TVTDLN 104

RESULT 28
US-10-116-273-8
; Sequence 8, Application US/10116273
; Publication No. US20020192183A1
; GENERAL INFORMATION:
; APPLICANT: Anne Dam Jensen
; TITLE OF INVENTION: Interferon gamma polypeptide variants
; FILE REFERENCE: 231us410 - INFG variants
; CURRENT APPLICATION NUMBER: US/10/116,273
; CURRENT FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentln Ver. 2.1
```

```
; SEQ ID NO 8
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: C-terminally
; OTHER INFORMATION: truncated [S99T]huIFNG
; -10-116-273-8

Query Match
3.2%; Score 6; DB 9; Length 136;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 TVTDLN 144
|||||
Db 99 TVTDLN 104

RESULT 29
US-10-116-273-7
; Sequence 7, Application US/10116273
; Publication No. US20020192183A1
; GENERAL INFORMATION:
; APPLICANT: Anne Dam Jensen
; TITLE OF INVENTION: Interferon gamma polypeptide variants
; FILE REFERENCE: 231us410 - INFG variants
; CURRENT APPLICATION NUMBER: US/10/116,273
; CURRENT FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 7
; LENGTH: 137
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: C-terminally
; OTHER INFORMATION: truncated [S99T]huIFNG
; -10-116-273-7

Query Match
3.2%; Score 6; DB 9; Length 137;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 TVTDLN 144
|||||
Db 99 TVTDLN 104

RESULT 30
US-10-116-273-6
; Sequence 6, Application US/10116273
; Publication No. US20020192183A1
; GENERAL INFORMATION:
; APPLICANT: Anne Dam Jensen
; TITLE OF INVENTION: Interferon gamma polypeptide variants
; FILE REFERENCE: 231us410 - INFG variants
; CURRENT APPLICATION NUMBER: US/10/116,273
; CURRENT FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 6
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: C-terminally
; OTHER INFORMATION: truncated [S99T]huIFNG
; -10-116-273-6

Query Match
3.2%; Score 6; DB 9; Length 138;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 139 TVTDLN 144  
| | | | |  
Db 99 TVTDLN 104

RESULT 31  
US-10-116-273-5  
; Sequence 5, Application US/10116273  
; Publication No. US20020192183A1  
; GENERAL INFORMATION:  
; APPLICANT: Anne Dam Jensen  
; TITLE OF INVENTION: Interferon gamma polypeptide variants  
; FILE REFERENCE: 231us410 - INFG variants  
; CURRENT APPLICATION NUMBER: US/10/116,273  
; CURRENT FILING DATE: 2002-04-04  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 139  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: C-terminally  
US-10-116-273-5

Query Match 3.2%; Score 6; DB 9; Length 139;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 139 TVTDLN 144  
| | | | |  
Db 99 TVTDLN 104

RESULT 32  
US-10-116-273-4  
; Sequence 4, Application US/10116273  
; Publication No. US20020192183A1  
; GENERAL INFORMATION:  
; APPLICANT: Anne Dam Jensen  
; TITLE OF INVENTION: Interferon gamma polypeptide variants  
; FILE REFERENCE: 231us410 - INFG variants  
; CURRENT APPLICATION NUMBER: US/10/116,273  
; CURRENT FILING DATE: 2002-04-04  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 140  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: C-terminally  
US-10-116-273-4

Query Match 3.2%; Score 6; DB 9; Length 140;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 139 TVTDLN 144  
| | | | |  
Db 99 TVTDLN 104

RESULT 33  
US-10-116-273-3  
; Sequence 3, Application US/10116273  
; Publication No. US20020192183A1  
; GENERAL INFORMATION:  
; APPLICANT: Anne Dam Jensen  
; TITLE OF INVENTION: Interferon gamma polypeptide variants  
; FILE REFERENCE: 231us410 - INFG variants

; CURRENT APPLICATION NUMBER: US/10/116,273  
; CURRENT FILING DATE: 2002-04-04  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 141  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: C-terminally  
US-10-116-273-3

Query Match 3.2%; Score 6; DB 9; Length 141;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 139 TVTDLN 144  
| | | | |  
Db 99 TVTDLN 104

RESULT 34  
US-10-116-273-2  
; Sequence 2, Application US/10116273  
; Publication No. US20020192183A1  
; GENERAL INFORMATION:  
; APPLICANT: Anne Dam Jensen  
; TITLE OF INVENTION: Interferon gamma polypeptide variants  
; FILE REFERENCE: 231us410 - INFG variants  
; CURRENT APPLICATION NUMBER: US/10/116,273  
; CURRENT FILING DATE: 2002-04-04  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 142  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: C-terminally  
US-10-116-273-2

Query Match 3.2%; Score 6; DB 9; Length 142;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 139 TVTDLN 144  
| | | | |  
Db 99 TVTDLN 104

RESULT 35  
US-09-308-207-60  
; Sequence 60, Application US/09308207  
; Publication No. US20030022323A1  
; GENERAL INFORMATION:  
; APPLICANT: MARIA DIAZ-TORRES ET AL.  
; TITLE OF INVENTION: METHOD FOR THE RECOMBINANT  
; PRODUCTION OF 1,3 PROPANEDIOL

NUMBER OF SEQUENCES: 68  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genencor International, Inc.  
STREET: 4 Cambridge Place  
CITY: Rochester  
STATE: NY  
COUNTRY: U.S.A  
ZIP: 14618  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows

```
SOFTWARE: FastSeq for Windows Version 2.0b
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/308,207
; FILING DATE: 13-May-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/030,601
; FILING DATE: 13-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Glaister, Debra
; REGISTRATION NUMBER: 33,888
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-864-7620
; TELEFAX: 650-845-6504
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 142 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. US2003002233A1e
; SEQUENCE DESCRIPTION: SEQ ID NO: 60:
US-09-308-207-60

Query Match      3.2%; Score 6; DB 9; Length 142;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 QAVEAK 37
Db 20 QAVEAK 25

RESULT 36
US-10-116-273-1
; Sequence 1, Application US/10116273
; Publication No. US20020192183A1
; GENERAL INFORMATION:
; APPLICANT: Anne Dam Jensen
; TITLE OF INVENTION: Interferon gamma polypeptide variants
; FILE REFERENCE: 231us410 - INFg variants
; CURRENT APPLICATION NUMBER: US/10/116,273
; FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: [S99T]huIFNG
US-10-116-273-1

Query Match      3.2%; Score 6; DB 9; Length 143;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 TVTDLN 144
Db 99 TVTDLN 104

RESULT 37
US-10-147-874-6
; Sequence 6, Application US/10147874
; Publication No. US20030027237A1
; GENERAL INFORMATION:
; APPLICANT: TARDIEUX, ISABELLE
; APPLICANT: DELORE, VIOLETTE
; TITLE OF INVENTION: SERINE-THREONINE PHOSPHATASE PROTEIN OF A PARASITIC ORGANISM OF T
; TITLE OF INVENTION: APICOMPLEXA PHYLLUM; APPLICATIONS IN THERAPEUTICS
```

```
; FILE REFERENCE: 223316US0
; CURRENT APPLICATION NUMBER: US/10/147,874
; CURRENT FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: US 60/291,609
; PRIOR FILING DATE: 2001-05-18
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-147-874-6

Query Match      3.2%; Score 6; DB 9; Length 151;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GEVAKL 45
Db 68 GEVAKL 73

RESULT 38
US-09-765-272-224
; Sequence 224, Application US/09765272
; Patent No. US20020061545A1
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/765,272
; FILING DATE: 22-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 224:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 224:
US-09-765-272-224

Query Match      3.2%; Score 6; DB 10; Length 153;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 LDKLIE 102
Db 145 LDKLIE 150
```

```
RESULT 39
US-09-864-761-48707
; Sequence 48707, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aomicca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 48707
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007480.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.94
; OTHER INFORMATION: EST HUMAN HIT: AA132806.1, EVALUATE 2.00e-10
; OTHER INFORMATION: SWISSPROT HIT: Q03112, EVALUATE 2.00e-69
US-09-864-761-48707
```

```
Query Match 3.2%; Score 6; DB 10; Length 160;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 40 GEVKKL 45
Db 88 GEVKKL 93
```

```
RESULT 40
US-10-101-464A-548
; Sequence 548, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 548
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-10-101-464A-548
```

```
Query Match 3.2%; Score 6; DB 9; Length 161;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 144 NLVPLI 149
Db 137 NLVPLI 142
```

```
Search completed: April 9, 2003, 14:29:18
Job time : 53.945 secs
```

01 Bacterifili: Proteobacteria; gamma subdivision; Pasteurellaceae;  
02 Haemophilus  
03  
04  
05  
06  
07  
08  
09  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
8

```

Db 2 YTIPELGYAVNALBEPHQAQIMEIHS 28

RESULT 40
AC SODM_CHAFE STANDARD; PRT; 224 AA.
AC 096347;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Superoxide dismutase, [mn], mitochondrial precursor (BC 1.15.1.1).
OS Charybdis feriatius (Crab).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;
OC Malacostraca; Eumalacostraca; Eucarida; Decapoda; Ploceymata;
OC Brachyura; Eubrachyura; Portunidea; Portunidae; Charybdis.
OX NCBI_TaxID=65693;
RN [1]
RP SEQUENCE FROM N.A.
RP Lin C.T., Lai Y.S., Kuo T.J., Chang T.C.;
RT "Molecular cloning, expression, and characterization of a cDNA
RT encoding Mn-superoxide dismutase from crab Charybdis feriatius.",
RL submitted (AUG-1997) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems.
CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -1- COFACTOR: Manganese (by similarity).
CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE
CC FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/sib-ch).
CC or send an email to license@sib-sib.ch).
-----
DR EMBL; AF019411; AAD01640.1; -.
DR HSSP; P04179; IABM.
DR InterPro; IPR001189; SODismutase.
DR Pfam; PF00081; sodfe, 1.
DR Pfam; PF02777; sodfe.C; 1.
DR ProDom; PD000475; SODismutase; 1.
DR PROSITE; PS00086; SOD_MN; 1.
KW Oxidoreductase; Manganese; Mitochondrion; Transit peptide.
KW TRANSIT 1 20 MITOCHONDRION (BY SIMILARITY).
FT CHAIN 21 224 SUPEROXIDE DISMUTASE [MN].
FT METAL 46 46 MANGANESE (BY SIMILARITY).
FT METAL 94 94 MANGANESE (BY SIMILARITY).
FT METAL 177 177 MANGANESE (BY SIMILARITY).
FT METAL 181 181 MANGANESE (BY SIMILARITY).
SQ SEQUENCE 224 AA; 24527 MW; 107CFL9382E9138A CRC64;

Query Match 46.2%; Score 67; DB 1; Length 224;
Best Local Similarity 39.3%; Pred. No. 0.0068;
Matches 11; Conservative 10; Mismatches 7; Indels 0; Gaps 0;

QY 1 KYSLPELDYFSGATPEPIYSGQINELIXYT 28
DB 21 KHTLPDLPEYDYGALPTISAEIMQHHS 48

Search completed: April 9, 2003, 13:57:53
Job time : 4.85321 secs

```

```

RESULT 37
SODM_XANCA STANDARD; PRT; 23 AA.
ID SODM_XANCA STANDARD; PRT; 23 AA.
AC P36215;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
OS Superoxide dismutase [Mn], mitochondrial (EC 1.15.1.1) (Fragment).
OC Rana catesbeiana (Bull. frog).
OC Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranidae; Rana.
OX NCBI_TaxID=8400;
RN [1]
RP SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=87126854; PubMed=3492965;
RA Abe Y., Okazaki T.;
" Purification and properties of the manganese superoxide dismutase
  from the liver of bullfrog, Rana catesbeiana.";
Arch. Biochem. Biophys. 253:241-248(1987).
CC -1- FUNCTION: Destroys radicals which are normally produced within the
  cells and which are toxic to biological systems.
CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -1- COFACTOR: Manganese.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE
  FAMILY.
DR InterPro: IPR001189; SODismutase.
DR Pfam: PF00081; sode; 1.
DR ProDom: PD000475; SODismutase; 1.
DR PROSITE: PS00088; SOD MN; PARTIAL.
KM Oxidoreductase; Manganese; Mitochondrion.
FT NON TER 23
FT SEQUENCE 23 AA; 2594 MW; 5D80ED9B0E04FE25 CRC64;
SQ
Query Match 46.2%; Score 67; DB 1; Length 23;
Best Local Similarity 50.0%; Pred. No. 0.00054;
Matches 11; Conservative 7; Mismatches 4; Indels 0; Gaps 0;
QY 1 KYSLPELDYFSATEPYISGOI 22
DB 1 KHTLPDLPLVDFGALQPHISAEI 22
RESULT 38
SODM_XANCP STANDARD; PRT; 203 AA.
ID SODM_XANCP STANDARD; PRT; 203 AA.
AC P53654;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
OS Superoxide dismutase [Mn] (EC 1.15.1.1).
GN SODA OR SOD OR SODM OR XCC2278.
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=8004;
RX MEDLINE=96405948; PubMed=8810073;
RA Smith S.G., Wilson T.J., Dow J.M., Daniels M.J.;
" A gene for superoxide dismutase from Xanthomonas campestris pv.
  campestris and its expression during bacterial-plant interactions.";
Mol. Plant Microbe Interact. 9:584-593(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furian L.R.,

```

```

RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarote G., Canavari F., Cardozo J., Chamberg F., Clapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidams J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spindola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
  host specificities.";
RL Nature 417:459-463(2002).
CC -1- FUNCTION: Destroys radicals which are normally produced within the
  cells and which are toxic to biological systems.
CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -1- COFACTOR: Manganese.
CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE
  FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL: U42464; AAB47971.1; -.
CC DR EMBL: AE012335; AAM41557.1; -.
CC DR HSSP: P00448; 1VEM.
CC DR InterPro: IPR001189; SODismutase.
CC DR Pfam: PF00081; sode; 1.
CC DR ProDom: PD000475; SODismutase; 1.
CC DR PROSITE: PS00088; SOD MN; 1.
CC KM Oxidoreductase; Manganese.
CC FT METAL 27
CC FT METAL 81
CC FT METAL 164
CC FT METAL 168
CC FT CONFLICT 53
CC FT CONFLICT 53
CC FT CONFLICT 88
CC FT CONFLICT 123
CC FT CONFLICT 141
CC FT CONFLICT 141
CC FT CONFLICT 180
CC FT CONFLICT 188
CC FT CONFLICT 192
CC FT CONFLICT 192
CC SQ SEQUENCE 203 AA; 22721 MW; C2C41P9FB150C860 CRC64;
Query Match 46.2%; Score 67; DB 1; Length 203;
Best Local Similarity 48.1%; Pred. No. 0.0061;
Matches 13; Conservative 4; Mismatches 10; Indels 0; Gaps 0;
QY 2 YSLPELDYFSATEPYISGOINIXT 28
DB 3 YTLPLQPLVAYDALPNDIAQTWEIHHT 29
RESULT 39
SODM_HABIN STANDARD; PRT; 214 AA.
ID SODM_HABIN STANDARD; PRT; 214 AA.
AC P43725;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
OS Superoxide dismutase [Mn] (EC 1.15.1.1).
GN SODA OR H11088.
OS Haemophilus influenzae.

```



```

RESULT 33
SODM_CAVPO STANDARD; PRT; 211 AA.
ID SODM_CAVPO
AC P49114;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Superoxide dismutase [Mn], mitochondrial precursor (EC 1.15.1.1)
DE (Fragment).
GN SOD2.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hartley; TISSUE=Lung;
RX MEDLINE=96180320; PubMed=8597602;
RA Yuan H.T., Bingle C.D., Kelly F.J.;
RA "Differential patterns of antioxidant enzyme mRNA expression in
RA guinea pig lung and liver during development.";
RA Biochim. Biophys. Acta 1305:163-171(1996).
CC -1- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems.
CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -1- COFACTOR: Manganese (By similarity).
CC -1- SUBUNIT: HOMOTETRAMER (By similarity).
CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U39843; AAC52719.1; -.
CC HSSP; P04179; IABM.
CC InterPro; IPR001189; SODismutase.
CC Pfam; PF00061; Sodfe; 1.
CC DR Pfam; PF02777; Sodfe; C; 1.
CC DR ProDom; PD000475; SODismutase; 1.
CC DR PROSITE; PS00088; SOD_MN; 1.
CC KM Oxidoreductase; Manganese; Mitochondrion; Transit peptide.
CC TRANSIT 1 24 MITOCHONDRION (By similarity).
CC CHAIN 25 >211 SUPEROXIDE DISMUTASE [Mn].
CC METAL .50 50 MANGANESE (By similarity).
CC FT METAL 98 98 MANGANESE (By similarity).
CC FT METAL 183 183 MANGANESE (By similarity).
CC FT METAL 187 187 MANGANESE (By similarity).
CC FT NON_TER 211 211
CC SQ SEQUENCE 211 AA; 23235 MW; 2032559161993E2D CRC64;
Query Match 46.3%; Score 68; DB 1; Length 211;
Best Local Similarity 35.7%; Pred. No. 0.0045;
Matches 10; Conservative 12; Mismatches 6; Indels 0; Gaps 0;
QY 1 KYSLPELDYEPSPATEPIYSGQINEIXYT 28
Db 25 KHSLELDLPDYDALPHINAEIMQLHHS 52
RESULT 34
SODF_STRCO STANDARD; PRT; 212 AA.
ID SODF_STRCO
AC 051917;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Superoxide dismutase [Fe-Zn] 1 (EC 1.15.1.1) (FesOD 1) (SOD2).

```

```

GN SODF1 OR SODF OR SODB OR SCO2633 OR SC984A.03.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Streptomyces; Streptomycesaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 10147;
RX MEDLINE=98215165; PubMed=9555880;
RA Kim E.-J., Chung H.-J., Suh B., Hah Y.C., Roe J.-H.;
RT "Expression and regulation of the sodF gene encoding iron- and zinc-
RT containing superoxide dismutase in Streptomyces coelicolor Muller.";
RL J. Bacteriol. 180:2014-2020(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=9250253; PubMed=10231572;
RA Chung H.J., Kim E.J., Suh B., Choi J.H., Roe J.H.;
RT "Duplicate genes for Fe-containing superoxide dismutase in
RT Streptomyces coelicolor A3(2).";
RL Gene 231:87-93(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieffer H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieffer T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rebbinowitsch E., Rajendram M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorek A., Woodward J., Barrett B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
RN [4]
RP SEQUENCE OF 1-15.
RC STRAIN=ATCC 10147;
RX MEDLINE=97054607; PubMed=8898904;
RA Kim F.-J., Kim H.-P., Han Y.V., Roe J.-H.;
RT "Differential expression of superoxide dismutases containing Ni and
RT Fe/Zn in Streptomyces coelicolor.";
RL Eur. J. Biochem. 241:178-185(1996).
CC -1- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems.
CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -1- COFACTOR: Iron and zinc.
CC -1- SUBUNIT: TETRAMER (POTENTIAL).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -1- INDUCTION: INHIBITED BY NICKEL.
CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF012087; AAC46274.1; -.
CC DR EMBL; AF099014; AAD33128.1; -.
CC DR EMBL; AL138662; CAB71808.1; -.
CC DR HSSP; P80293; IAVM.
CC DR InterPro; IPR001189; SODismutase.
CC DR Pfam; PF00081; Sodfe; 1.
CC DR Pfam; PF02777; Sodfe; C; 1.
CC DR ProDom; PD000475; SODismutase; 1.
CC DR PROSITE; PS00088; SOD_MN; 1.
CC KM Oxidoreductase; Iron; Zinc; Complete proteome.

```



FT METAL 189 189 MANGANESE (BY SIMILARITY)  
 FT METAL 193 193 MANGANESE (BY SIMILARITY)  
 SO SEQUENCE 228 AA; 25504 MW; 9F4D75B2ADADDF6 CRC64;

Query Match 47.6%; Score 69; DB 1; Length 228;  
 Best Local Similarity 46.2%; Pred. No. 0.0034;  
 Matches 12; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 2 XSLPELDYEFSAPEYISQINEIXY 27  
 DB 28 FSLPDLPYDALEPAISGIMOLHH 53

## RESULT 31

SODM AGABI STANDARD; PRT; 200 AA.

AC Q9P4T6;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 Superoxide dismutase [Mn], mitochondrial precursor (EC 1.15.1.1).  
 SOD.

OS Agaricus bisporus (Common mushroom).  
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;  
 OC Agaricales; Agaricaceae; Agaricus.  
 OX NCBI\_TaxId=5341;  
 RX [1]

RP SEQUENCE FROM N.A.

RA STRAIN=Horet U3;  
 RC Eastwood D.C., Bains N.K., Henderson J., Burton K.S.;

RT "Oxidative stress in the harvested mushroom, Agaricus bisporus."  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: Destroys radicals which are normally produced within the  
 cells and which are toxic to biological systems.  
 CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).  
 CC -1- COFACTOR: Manganese (by similarity).  
 CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.  
 CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE  
 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AJ04469; CAB94731.1; -

CC HSSP; P04179; 1VAR.

CC InterPro; IPR001189; SODismutase.

CC Pfam; PF00081; sode1.1.

CC Pfam; PF02777; sode1.1.

CC ProDom; PD000475; SODismutase; 2.

CC PROSITE; PS00088; SOD\_MN; 1.

CC Oxidoreductase; Manganese; Mitochondrion; Transit peptide.

CC TRANSIT 1 ? MITOCHONDRION (POTENTIAL).

CC CHAIN 1 ? SUPEROXIDE DISMUTASE [MN].

CC METAL 27 27 MANGANESE (BY SIMILARITY).

CC METAL 72 72 MANGANESE (BY SIMILARITY).

CC METAL 157 157 MANGANESE (BY SIMILARITY).

CC METAL 161 161 MANGANESE (BY SIMILARITY).

CC SEQUENCE 200 AA; 22194 MW; 9758B1DD1F64FF19 CRC64;

Query Match 46.9%; Score 68; DB 1; Length 200;  
 Best Local Similarity 54.2%; Pred. No. 0.0042;  
 Matches 13; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 4 LPELDYEFSAPEYISQINEIXY 27  
 DB 5 LPELDYEFSAPEYISQINEIXY 28

RESULT 32  
 SODF SULAC STANDARD; PRT; 210 AA.

AC 008713;

DT 01-OCT-1994 (Rel. 30, Created)

DT 01-OCT-1994 (Rel. 30, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Superoxide dismutase [Fe] (EC 1.15.1.1).

GN SOD

OS Sulfolobus acidocaldarius.

OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;

OC Sulfolobus.

OX NCBI\_TaxId=2285;

RP SEQUENCE FROM N.A.

RA MEDLINE=93326644; PubMed=8334170;

RA Klenk H.-P., Schleper C., Schwass V., Brudler R.;

RT "Nucleotide sequence, transcription and phylogeny of the gene

RL encoding the superoxide dismutase of Sulfolobus acidocaldarius.";

RL Biochim. Biophys. Acta 1174:95-98 (1993).

RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).

RC STRAIN=ATCC 33909 / NCIB 11770 / DSM 639;

RX MEDLINE=99069655; PubMed=9878438;

RA Knapp S., Kardinahl S., Helgren N., Tibbelin G., Schaefer G.,

RA Ladenstein R.;

RT "Refined crystal structure of a superoxide dismutase from the

RT hyperthermophilic archaeon Sulfolobus acidocaldarius at 2.2-A

RT resolution.";

RL J. Mol. Biol. 285:689-702 (1999).

CC -1- FUNCTION: Destroys radicals which are normally produced within the

CC cells and which are toxic to biological systems.

CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).

CC -1- COFACTOR: IRON.

CC -1- SUBUNIT: HOMOTETRAMER AT HIGH TEMPERATURE; HOMODIMER AT ROOM

CC TEMPERATURE.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic.

CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE

CC FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; X63386; CAA44993.1; -

CC PIR; S34616; S34616.

CC PDB; 1B06; 18-NOV-99.

CC InterPro; IPR001189; SODismutase.

CC Pfam; PF00081; sode1.1.

CC Pfam; PF02777; sode1.1.

CC ProDom; PD000475; SODismutase; 1.

CC PROSITE; PS00088; SOD\_MN; FALSE NEG.

CC Oxidoreductase; Iron; 3D-structure.

CC INIT\_MET 0 0

CC METAL 33 33 IRON.

CC METAL 84 84 IRON.

CC METAL 170 170 IRON.

CC METAL 174 174 IRON.

CC SEQUENCE 210 AA; 24135 MW; 086CAB277D99FBB CRC64;

Query Match 46.9%; Score 68; DB 1; Length 210;  
 Best Local Similarity 44.4%; Pred. No. 0.0044;  
 Matches 12; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 1 XSLPELDYEFSAPEYISQINEIXY 27  
 DB 8 XSLPELDYEFSAPEYISQINEIXY 34

GN SODF.  
 OS Bacillus subtilis.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=1423;  
 RN [1]  
 RA SEQUENCE FROM N.A.  
 RP Lapidus A., Galleron N., Sorokin A., Ehrlich S.D.  
 RT "Sequence analysis of the Bacillus subtilis chromosome region between  
 the t<sub>er</sub>C and o<sub>da</sub>B loci cloned in a yeast artificial chromosome."  
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=9804403; PubMed=9384377;  
 RA Kunze F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
 Azevedo V., Berteiro M.G., Bessieres P., Bolyon A., Borchert S.,  
 Borries R., Boutsier L., Brans A., Brann M., Bridgell S.C., Bron S.,  
 Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,  
 Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,  
 Denicot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,  
 Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
 Fritz C.Y., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
 Ghim S.Y., Glaeser P., Goffeau A., Golligly E.J., Grandi G.,  
 Guiseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,  
 Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
 Joris B., Karamata D., Kasehara Y., Kleier-Blanchard M., Klein C.,  
 Kobayashi Y., Koelter P., Koningsstein G., Krogh S., Kumano M.,  
 Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
 Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,  
 Medina D., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
 Neone D., O'Reilly M., Ogawa K., Ogiwara A., Oudga B., Park S.H.,  
 Paro V., Pohl T.M., Portetelle D., Portwille S., Prescott A.M.,  
 Pressac E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
 Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadleir Y.,  
 Sato T., Scanlan E., Schleich S., Schroeter R., Scofield F.,  
 Sekiuchi J., Sekowska A., Serró S.J., Serró P., Shin B.S., Solito B.,  
 Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
 Takenaka M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,  
 Tosa V., Uchiyama S., Vandenbol M., Vanlier P., Vassarotti A.,  
 Vardi A., Wambutt R., Wedler E., Weller H., Welternecker T.,  
 Winters P., Wipet A., Yamamoto H., Yamane K., Yano K., Yara K.,  
 Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,  
 RA "The complete genome sequence of the Gram-positive bacterium Bacillus  
 subtilis."  
 RT Nature 390:249-256(1997).  
 RL  
 CC -1- FUNCTION: Destroys radicals which are normally produced within the  
 cells and which are toxic to biological systems  
 CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).  
 CC -1- COFACTOR: Iron (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE  
 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: AF027868; AAB84442.1; -  
 CC EMBL: Z99114; CAB13825.1; -  
 CC HSSP: P80293; IAVM.  
 CC Subtilist; BG12676; sodF.  
 CC InterPro: IPR001189; SODismutase.  
 CC Pfam: PF00081; sodF; 1.  
 CC Pfam: PF02777; sodF; 1.  
 CC ProDom: PD000475; SODismutase; 1.  
 CC PROSITE: PS00088; SOD MN; 1.  
 CC Oxidoreductase; Iron; Complete proteome.  
 KW METAL 104 104 IRON (BY SIMILARITY).  
 FT METAL 152 152 IRON (BY SIMILARITY).  
 FT METAL 236 236 IRON (BY SIMILARITY).

FT METAL 240 240 IRON (BY SIMILARITY).  
 SQ SEQUENCE 281 AA; 33477 MW; 7F36AC0460E74DB0 CRC64;  
 Query Match 48.3%; Score 70; DB 1; Length 281;  
 Best Local Similarity 50.0%; Pred. No. 0.0031;  
 Matches 14; Conservative 6; Mismatches 8; Indels 0; Gaps 0;  
 Oy 1 KYSLEPDIYEFSPATPYISGQINEIXYT 28  
 Db 79 KHVLPKLPYKXSALPYISRDIMLHHT 106  
 RESULT 30  
 SODM\_NICEPL STANDARD; PRT; 228 AA.  
 ID SODM\_NICEPL  
 AC P11796;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Superoxide dismutase [Mn], mitochondrial precursor (EC 1.15.1.1).  
 GN SODA.  
 OS Nicotiana glauca (leadwort-leaved tobacco).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.  
 OX NCBI\_TaxID=4092;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=P2;  
 RX MEDLINE=8921633; PubMed=2540959;  
 RA Bowler C., Alliotte T., de Loose M., van Montagu M., Inze D.,  
 RT "The induction of manganese superoxide dismutase in response to  
 stress in Nicotiana glauca.";  
 RL EMBO J. 8:31-38(1989).  
 RN [2]  
 RP SEQUENCE OF 25-47.  
 RA Bauw G., de Loose M., Inze D., van Montagu M., Vandekerckhove J.,  
 RT "Alterations in the phenotype of plant cells studied by NH<sub>2</sub>-terminal  
 RT amino acid-sequence analysis of proteins electrophoretically  
 RT separated from total extracts."  
 RL Proc. Natl. Acad. Sci. U.S.A. 84:4806-4810(1987).  
 CC -1- FUNCTION: Destroys radicals which are normally produced within the  
 cells and which are toxic to biological systems.  
 CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).  
 CC -1- COFACTOR: Manganese.  
 CC -1- SUBUNIT: HOMOTETRAMER.  
 CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.  
 CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE  
 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: X14482; CAA22643.1; -  
 CC EMBL: A05355; CAA00450.1; -  
 CC PIR: S03639; S03639.  
 CC PIR: I28027; I28027.  
 CC HSSP: P04179; IAP6.  
 CC InterPro: IPR001189; SODismutase.  
 CC Pfam: PF00081; sodF; 1.  
 CC Pfam: PF02777; sodF; 1.  
 CC ProDom: PD000475; SODismutase; 1.  
 CC PROSITE: PS00088; SOD MN; 1.  
 CC Oxidoreductase; Manganese; Mitochondrion;  
 KW TRANSIT 1 24  
 FT CHAIN 25 228 SUPEROXIDE DISMUTASE [MN].  
 FT METAL 52 52 MANGANESE (BY SIMILARITY).  
 FT METAL 100 100 MANGANESE (BY SIMILARITY).

```

RA Creasy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,
RA P. G., Miltner J., Sellers P., Gill J.E., Feldlyum T.V.,
RA Preus D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,
RA Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asanizu E.,
RA Sasamoto S., Kimura T., Ideawara K., Kawashima K., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakayama S., Nakazaki N., Shino S., Takeuchi C., Wada T.,
RA Matanabe A., Yamada M., Yasuda M., Tabata S.;
RA "Sequence and analysis of chromosome 3 of the plant Arabidopsis
RT thaliana";
RL Nature 408:820-822(2000).
RL [3]
RN SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT "Riken Arabidopsis full length cDNA clones (RAFLs) sequenced by the
RT SGP consortium (Salk/Stanford/PGSC).";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems.
CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -1- COFACTOR: Manganese (By similarity).
CC -1- SUBUNIT: HOMOTETRAMER (By similarity).
CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF061518; AAC24832.1; -.
CC DR EMBL: AC009991; AAF01529.1; -.
CC DR EMBL: AY072495; AAL66910.1; -.
CC DR EMBL: AY059807; AAL24289.1; -.
CC DR HSSP: P04179; IABM.
CC DR InterPro: IPR001189; SODismutase.
CC DR Pfam: PF00081; sode, 1.
CC DR Pfam: PF02777; sode_C, 1.
CC DR ProDom: PD000475; SODismutase; 1.
CC DR PROSITE: PS00088; SOD_MN; 1.
CC KM Oxidoreductase; Manganese; Mitochondrion; Transit peptide.
CC FT TRANSIT 1 29 MITOCHONDRION (By similarity).
CC FT CHAIN 30 231 SUPEROXIDE DISMUTASE [MN].
CC FT METAL 59 59 MANGANESE (By similarity).
CC FT METAL 103 103 MANGANESE (By similarity).
CC FT METAL 192 192 MANGANESE (By similarity).
CC FT METAL 196 196 MANGANESE (By similarity).
CC FT CONFLICT 169 169 V -> F (IN REF. 1).
CC SQ SEQUENCE 231 AA; 25444 MW; 2DBD5560A9EBAD7D CRC64;
CC -----
Query Match 48.3%; Score 70; DB 1; Length 231;
Best Local Similarity 46.2%; Pred. No. 0.0025;
Matches 12; Conservative 8; Mismatches 6; Indels 0; Gaps 0;
QY 2 YSLPELDYFSGATEPYISGOINEIXY 27
Db 31 FTLPDLVYDYGALBPALISGEIMQIH 56

```

```

OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Euroside I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.
OC NCBI_TaxId=3888;
RN [1]
RN SEQUENCE FROM N.A.
RP TISSUE=Leaf;
RX MEDLINE=92032795; PubMed=1932701;
RA Wong-Vega L., Burke J.J., Allen R.D.;
RT "Isolation and sequence analysis of a cDNA that encodes pea manganese
RT superoxide dismutase.";
RL Plant Mol. Biol. 17:1271-1274(1991).
RN [2]
RN REVISIONS.
RC STRAIN=cv. Alaska; TISSUE=Isolated bud;
RA Jaradat T., Wong-Vega L., Allen R.D.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems.
CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -1- COFACTOR: Manganese (By similarity).
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X60170; CAA42737.1; -.
CC DR EMBL: U30841; AAA74442.1; -.
CC DR PIR: S18343; DSPMN.
CC DR HSSP: P04179; IABM.
CC DR InterPro: IPR001189; SODismutase.
CC DR Pfam: PF00081; sode, 1.
CC DR Pfam: PF02777; sode_C, 1.
CC DR ProDom: PD000475; SODismutase; 1.
CC DR PROSITE: PS00088; SOD_MN; 1.
CC KM Oxidoreductase; Manganese; Mitochondrion; Transit peptide.
CC FT TRANSIT 1 36 MITOCHONDRION (PROBABLE).
CC FT CHAIN 37 233 SUPEROXIDE DISMUTASE [MN].
CC FT METAL 64 64 MANGANESE (By similarity).
CC FT METAL 112 112 MANGANESE (By similarity).
CC FT METAL 194 194 MANGANESE (By similarity).
CC FT METAL 198 198 MANGANESE (By similarity).
CC FT CONFLICT 160 160 Q -> QASGVV (IN REF. 1).
CC FT CONFLICT 176 176 T -> TAN (IN REF. 1).
CC FT CONFLICT 192 192 G -> W (IN REF. 1).
CC SQ SEQUENCE 233 AA; 25822 MW; DB246D1FF0AFAFF9 CRC64;
CC -----
Query Match 48.3%; Score 70; DB 1; Length 233;
Best Local Similarity 46.2%; Pred. No. 0.0025;
Matches 12; Conservative 8; Mismatches 6; Indels 0; Gaps 0;
QY 2 YSLPELDYFSGATEPYISGOINEIXY 27
Db 40 FTLPDLVYDYGALBPALISGEIMQIH 65

```

```

RESULT 28
SODM_PEA STANDARD; PRT; 233 AA.
AC P27084; Q43069;
DT 01-AUG-1992 (Rel. 23, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Superoxide dismutase [Mn], mitochondrial precursor [EC 1.15.1.1].
GN SODA.

```

```

RESULT 29
SODF_BACSU STANDARD; PRT; 281 AA.
AC Q35023;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable superoxide dismutase [Fe] (EC 1.15.1.1).

```

DR PRODOM: PD000475; SODismutase; 1.  
 DR PROSITE; PS00088; SOD\_MN; 1.  
 KM Oxidoreductase; Iron; Complete proteome.  
 FT METAL 26 IRON (BY SIMILARITY).  
 FT METAL 81 IRON (BY SIMILARITY).  
 FT METAL 164 164 IRON (BY SIMILARITY).  
 FT METAL 168 168 IRON (BY SIMILARITY).  
 SQ SEQUENCE 203 AA; 23523 MW; ABA3A24EAC591716 CRC64;  
 Query Match 48.3%; Score 70; DB 1; Length 203;  
 Best Local Similarity 44.4%; Pred. No. 0.0021;  
 Matches 13; Conservative 6; Mismatches 9; Indels 0; Gaps 0;  
 QY 2 YSPLELDYFSATEPYISGQINEIYTT 28  
 Db 2 FKLPELDGYDAVEPYIDAKTMEIHHS 28  
 RESULT 26  
 PRODM\_CHLPN STANDARD; PRT; 207 AA.  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 15-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Superoxide dismutase (Mn) (EC 1.15.1.1).  
 GN SODA OR SODM OR CPN0057 OR CP0718.  
 OS Chlamydia pneumoniae (Chlamydia pneumoniae).  
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.   
 NCBI\_TaxId=83556;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AR39;  
 RX MEDLINE=20150255; PubMed=10192388;  
 RA Kilmann S., Mitchell W., Marache R., Lammell C., Fan J., Hyman R.W.,  
 RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.,  
 RA "Comparative genomes of Chlamydia pneumoniae and C. trachomatis";  
 RA Nat. Genet. 21:385-389(1999).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AR39;  
 RX MEDLINE=20150255; PubMed=10684935;  
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heideberg J.F.,  
 RA White O., Hickey E.K., Peterson J., Uterback T., Berry K., Bass S.,  
 RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,  
 RA Gilm M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,  
 RA Eisen J., Fraser C.M.;  
 RA "Genome sequences of Chlamydia trachomatis Morn and Chlamydia  
 pneumoniae AR39";  
 RA Nucleic Acids Res. 28:1397-1406(2000).  
 [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=J138;  
 RX MEDLINE=20330349; PubMed=10871362;  
 RA Shiba T., Ishii K., Hattori M., Kunara S., Nakazawa T.,  
 RA "Comparison of whole genome sequences of Chlamydia pneumoniae J138  
 from Japan and CML029 from USA";  
 RA Nucleic Acids Res. 28:2311-2314(2000).  
 RL  
 CC -1- FUNCTION: Destroy radicals which are normally produced within the  
 CC cells and which are toxic to biological systems.  
 CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).  
 CC -1- COFACTOR: Manganese (By similarity).  
 CC -1- SUBUNIT: HOMODIMER (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE  
 CC FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 CC -----  
 DR EMBL; AE001591; AAD18210.1; -.  
 DR EMBL; AE002230; AAF38524.1; -.  
 DR EMBL; AF002545; BAA98268.1; -.  
 DR HSSP; P04179; IAP6.  
 DR PHCI-2DPAGE; Q929C4; -.  
 DR TIGR; CP0718; -.  
 DR InterPro; IPR001189; SODismutase.  
 DR Pfam; PF00081; sode; 1.  
 DR Pfam; PF02777; sode\_C; 1.  
 DR PRODOM: PD000475; SODismutase; 1.  
 DR PROSITE; PS00088; SOD\_MN; 1.  
 KM Oxidoreductase; Manganese; Complete proteome.  
 FT METAL 31 31 MANGANESE (BY SIMILARITY).  
 FT METAL 78 78 MANGANESE (BY SIMILARITY).  
 FT METAL 166 166 MANGANESE (BY SIMILARITY).  
 FT METAL 170 170 MANGANESE (BY SIMILARITY).  
 SQ SEQUENCE 207 AA; 23541 MW; 45A4A6FF096F8934 CRC64;  
 Query Match 48.3%; Score 70; DB 1; Length 207;  
 Best Local Similarity 61.9%; Pred. No. 0.0022;  
 Matches 13; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
 QY 2 YSPLELDYFSATEPYISGQI 22  
 Db 6 YSPLELDYDALEPVISSEI 26  
 RESULT 27  
 SODM\_ARATH STANDARD; PRT; 231 AA.  
 ID SODM\_ARATH  
 AC 081235; Q9SRK3;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DE Superoxide dismutase (Mn), mitochondrial precursor (EC 1.15.1.1).  
 GN SODA OR MSD1 OR ATG310920 OR P9F8.26.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.  
 NCBI\_TaxId=3702;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=98440686; PubMed=9765550;  
 RA Kliebenstein D.J., Monde R.A., Last R.L.;  
 RA "Superoxide dismutase in Arabidopsis: an eclectic enzyme family with  
 RA disparate regulation and protein localization";  
 RA Plant Physiol. 118:637-650(1998).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=21016720; PubMed=11130713;  
 RA Salasoulat M., Lemcke K., Rieger M., Ansege W., Unseid M.,  
 RA Fartman B., Valle G., Bloeker H., Perez-Alonso M., Obermaier B.,  
 RA Delgeny M., Boutry M., Griwell L.A., Mache R., Puigdomenech P.,  
 RA De Simone V., Choise N., Artiguenave F., Robert C., Brotier P.,  
 RA Wincker P., Cactolico L., Weissbach J., Saurin W., Quetier F.,  
 RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Bene V.,  
 RA Wurmback E., Drzonek H., Erle H., Jordan N., Bangert S.,  
 RA Wiedemann R., Kranz H., Voes H., Holland R., Brangert P.,  
 RA Vezzi A., D'Angelo M., Pallavicini A., Toppi S., Simonati B.,  
 RA Conrad A., Hornischer K., Kauer G., Loehner T.H., Nordstiek G.,  
 RA Reichelt J., Schaefer M., Schoen O., Bargues M., Terol J., Clement J.,  
 RA Navarro P., Collado C., Perez-Perez A., Outenwalder B., Duchemin D.,  
 RA Cooke R., Laude M., Berger-lilauro C., Purnelle B., Masny D.,  
 RA de Haan M., Maiores A.C., Alcaraz J.-P., Cortet A., Casasuberta E.,  
 RA Montfort A., Argitiro A., Flores M., Lignori R., Vitale D.,  
 RA Manhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Wewes H.-W.,  
 RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,  
 RA Rooney T., Rizzo J., Walts A., Uterback T., Fujii C.Y., Shea T.P.,





KW Oxidoreductase; Manganese; Mitochondrion; Transil peptide.  
 FT TRANSIT 1 24 MITOCHONDRION.  
 FT CHAIN 25 222 SUPEROXIDE DISMUTASE [MN].  
 FT METAL 50 50 MANGANESE (BY SIMILARITY).  
 FT METAL 98 98 MANGANESE (BY SIMILARITY).  
 FT METAL 183 183 MANGANESE (BY SIMILARITY).  
 FT METAL 187 187 MANGANESE (BY SIMILARITY).  
 FT CONFLICT 167 167 Q -> H (IN REF. 1).  
 SQ SEQUENCE 222 AA; 24674 MW; 8CCCE0E857B3138 CRC64;

Query Match 51.0%; Score 74; DB 1; Length 222;  
 Best Local Similarity 42.9%; Pred. No. 0.00056;  
 Matches 12; Conservative 10; Mismatches 6; Indels 0; Gaps 0;

1 KYSLPELDYFESATEPYISQINEIXY 28  
 DB 25 KHSLEPLDYFALPEPHINQIMQIHHHS 52

RESULT 20  
 SODF\_SULSO STANDARD; PRT; 210 AA.  
 AC P80857;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Superoxide dismutase [Fe] (EC 1.15.1.1).  
 GN SOD OR SSO0316.  
 OS Sulfolobus solfataricus.  
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;  
 OC Sulfolobus.  
 OC NCBI\_TaxId=2287;  
 RN [1]  
 RP SEQUENCE  
 RC STRAIN=ATCC 49255 / DSM 5833 / MT-4;  
 RX MEDLINE=98088931; Pubmed=9428655;  
 RA DeLio Russo A., Rullo R., Niciti G., Masullo M., Bocchini V.;  
 RT "Iron superoxide dismutase from the archaeon Sulfolobus solfataricus:  
 average hydrophobicity and amino acid weight are involved in the  
 adaptation of proteins to extreme environments.";  
 RL Biochim. Biophys. Acta 1343:23-30(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=9908843; Pubmed=980816;  
 RA Yamano S., Maruyama T.;  
 RT "An azide-insensitive superoxide dismutase from a hyperthermophilic  
 archaeon, Sulfolobus solfataricus.";  
 RL J. Biochem. 125:186-193(1999).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=ATCC 49255 / DSM 5833 / MT-4;  
 RX MEDLINE=21145482; Pubmed=11248699;  
 RA De Vendittis E., Ursby T., Rullo R., Gogliettino M.A., Masullo M.,  
 Bocchini V.;  
 RT "Phenylmethanesulfonyl fluoride inactivates an archaeal superoxide  
 dismutase by chemical modification of a specific tyrosine residue:  
 cloning, sequencing and expression of the gene coding for Sulfolobus  
 solfataricus superoxide dismutase.";  
 RL Eur. J. Biochem. 268:1794-1801(2001).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=ATCC 35092 / DSM 1617 / P2;  
 RX MEDLINE=21332296; Pubmed=11427726;  
 RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,  
 Aweyer M.J., Chan-Weiner C.C.-Y., Clausen I.G., Curtis B.A.,  
 De Moers A., Eranuso G., Fletcher C., Gordon P.M.K.,  
 Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,  
 Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,  
 Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,  
 Garrett R.A., Regan M.A., Sensen C.W., Van der Oost J.;  
 RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).  
 RN [5]

RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).  
 RX MEDLINE=99134398; Pubmed=9911259;  
 RA Ursby T., Adinolfi B.S., Al-Karadaghi S., de Vendittis E.,  
 Bocchini V.;  
 RT "Iron superoxide dismutase from the archaeon Sulfolobus solfataricus:  
 analysis of structure and thermostability.";  
 RL J. Mol. Biol. 286:189-205(1999).  
 CC -1- FUNCTION: Destroys radicals which are normally produced within the  
 cells and which are toxic to biological systems;  
 CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).  
 CC -1- COFACTOR: IRON.  
 CC -1- SUBUNIT: HOMOTETRAMER.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE  
 FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AB012620; BAA75509.1; -.  
 CC EMBL; Y15326; CAA75583.1; -.  
 CC EMBL; AE006666; AAK40652.1; -.  
 CC PDB; 1SS8; 09-APR-99.  
 CC InterPro; IPR001189; SODismutase.  
 CC Pfam; PF00081; sode, 1.  
 CC Pfam; PF02777; sode\_C, 1.  
 CC ProDom; PD000475; SODismutase; 1.  
 CC PROSITE; PS00088; SOD MN; FALSE NEG.  
 KW Oxidoreductase; Iron; 3D-structure; Complete proteome.  
 FT INIT MET 0  
 FT METAL 37 37 IRON.  
 FT METAL 84 84 IRON.  
 FT METAL 170 170 IRON.  
 FT METAL 174 174 IRON.  
 SQ SEQUENCE 210 AA; 24112 MW; 7918CF1292BF98B6 CRC64;

Query Match 50.3%; Score 73; DB 1; Length 210;  
 Best Local Similarity 51.9%; Pred. No. 0.00078;  
 Matches 14; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

1 KYSLPELDYFESATEPYISQINEIXY 27  
 DB 8 KYSLPELDYFESATEPYISQINEIXY 34

RESULT 21  
 SODF\_PYRAE STANDARD; PRT; 211 AA.  
 ID SODF\_PYRAE  
 AC G93724;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Superoxide dismutase [Fe] (EC 1.15.1.1).  
 GN SOD OR PAE0274.  
 OS Pyrobaculum aerophilum.  
 OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;  
 OC Thermoproteaceae; Pyrobaculum.  
 OC NCBI\_TaxId=13773;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=IM2 / ATCC 51768 / DSM 7523;  
 RA Baikalov C.J., Slupska M.M., Miller J.H.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBD databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=IM2 / ATCC 51768 / DSM 7523;  
 RX Pubmed=11792869;  
 RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,







RP SEQUENCE OF 25-39.  
 RC TISSUE=Heart;  
 RX MEDLINE=95203287; Pubmed=7895732;  
 RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;  
 RT "The human myocardial two-dimensional gel protein database: update  
 1994.";  
 RL Electrophoresis 15:1459-1465(1994).  
 RN [10]  
 RP SEQUENCE OF 25-39.  
 RC TISSUE=Heart;  
 RX MEDLINE=9607936; Pubmed=7498159;  
 RA Kovalyov L.I., Shishkin S.S., Etimochkin A.S., Kovalyova M.A.,  
 RA Kshova E.S., Egorov T.A., Msaalamov A.K.;  
 RT "The major protein expression profile and two-dimensional protein  
 database of human heart.";  
 RL Electrophoresis 16:1160-1169(1995).  
 RN [11]  
 RP SEQUENCE OF 25-39.  
 RC TISSUE=Breast carcinoma;  
 RX MEDLINE=9729304; Pubmed=9150946;  
 RA Rasmussen R.K., Ji H., Eddes J.S., Moritz R.L., Reid G.E.,  
 RA Simpson R.J., Dorow D.S.;  
 RT "Two-dimensional electrophoretic analysis of human breast carcinoma  
 proteins: mapping of proteins that bind to the SH3 domain of mixed  
 lineage kinase MLK2.";  
 RL Electrophoresis 18:588-598(1997).  
 RN [12]  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
 RX MEDLINE=93008233; Pubmed=1394426;  
 RA Borgstahl G.E.O., Parge H.E., Hickey M.J., Beyer W.F. Jr.,  
 RA Hallewell R.A., Tainer J.A.;  
 RT "The structure of human mitochondrial manganese superoxide dismutase  
 reveals a novel tetrameric interface of two 4-helix bundles.";  
 RL Cell 71:107-116(1992).  
 RN [13]  
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF VARIANT THR-82.  
 RX MEDLINE=96183289; Pubmed=8605177;  
 RA Borgstahl G.E.O., Parge H.E., Hickey M.J., Johnson M.J.,  
 RA Boissacq M., Hallewell R.A., Lepock J.R., Cabelli D.E.,  
 RA Tainer J.A.;  
 RT "Human mitochondrial manganese superoxide dismutase polymorphic  
 variant I158Tn reduces activity by destabilizing the tetrameric  
 interface.";  
 RL Biochemistry 35:4287-4297(1996).  
 RN [14]  
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF VARIANT ASN-167.  
 RX MEDLINE=98206886; Pubmed=9537988;  
 RA Heich Y., Guan Y., Tu C., Bratt P.J., Angerhofer A., Lepock J.R.,  
 RA Hickey M.J., Tainer J.A., Nick H.S., Silverman D.N.;  
 RT "Probing the active site of human manganese superoxide dismutase: the  
 role of glutamine 143.";  
 RL Biochemistry 37:4731-4739(1998).  
 RN [15]  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF VARIANT TYR-58.  
 RX MEDLINE=98206887; Pubmed=9537987;  
 RA Guan Y., Hickey M.J., Borgstahl G.E.O., Hallewell R.A., Lepock J.R.,  
 RA O'Connor D., Heich Y., Nick H.S., Silverman D.N., Tainer J.A.;  
 RT "Crystal structure of Y34F mutant human mitochondrial manganese  
 superoxide dismutase and the functional role of tyrosine 34.";  
 RL Biochemistry 37:4722-4730(1998).  
 CC -1- FUNCTION: Destroys radicals which are normally produced within the  
 cells and which are toxic to biological systems.  
 CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).  
 CC -1- COFACTOR: Manganese.  
 CC -1- SUBUNIT: HOMOTETRAMER.  
 CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.  
 CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE  
 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X59445; CAA42066.1; -;  
 CC EMBL: Y00472; CAA68533.1; -;  
 CC EMBL: Y00985; CAA68791.1; -;  
 CC EMBL: X07834; CAA30687.1; -;  
 CC EMBL: M36693; AAA36622.1; -;  
 CC EMBL: X15132; CAA33228.1; -;  
 CC EMBL: X14322; CAA32502.1; -;  
 CC EMBL: S77127; AAD14248.1; ALT\_SEQ.  
 CC PIR: A27297; DSHUN.  
 CC PIR: S00356; S00356.  
 CC PIR: S02230; S02230.  
 CC PIR: S11756; S11756.  
 CC PIR: S13162; S13162.  
 CC PDB: 1MSD; 15-JUL-93.  
 CC PDB: 1AMW; 31-OCT-93.  
 CC PDB: 1ONW; 07-JAN-98.  
 CC PDB: 1VAR; 10-JUN-98.  
 CC PDB: 1AP5; 28-JAN-98.  
 CC PDB: 1AP6; 28-JAN-98.  
 CC SWISS-2DPAGE: P04179; HUMAN.  
 CC HSC-2DPAGE: P04179; HUMAN.  
 CC Siema-2DPAGE: P04179; -;  
 CC Genew; HGNC:11180; SOD2.  
 CC MIM; 147460; -;  
 CC InterPro: IPR001189; SODismutase.  
 CC Pfam: PF02771; Sodfe; 1.  
 CC Pfam: PF02777; Sodfe; 1.  
 CC Prodom: PD000475; SODismutase; 1.  
 CC PROSITE: PS00088; SOD\_MN; 1.  
 CC Oxidoreductase; Manganese; Mitochondrion; Transic peptide;  
 CC 3D-structure; Polymorphism.  
 CC TRANSIT 1 24  
 CC CHAIN 25 222  
 CC METAL 50 50  
 CC METAL 98 98  
 CC METAL 183 183  
 CC METAL 187 187  
 CC METAL 82 82  
 CC VARIANT  
 CC FT CONFLICT 14 14  
 CC FT CONFLICT 16 16  
 CC FT CONFLICT 65 65  
 CC FT CONFLICT 66 66  
 CC FT CONFLICT 112 112  
 CC FT CONFLICT 123 123  
 CC FT CONFLICT 133 133  
 CC FT CONFLICT 148 149  
 CC FT CONFLICT 155 155  
 CC FT CONFLICT 155 155  
 CC FT HELIX 35 38  
 CC FT TURN 39 41  
 CC FT TURN 44 52  
 CC FT HELIX 53 53  
 CC TURN 53 53  
 CC Query Match 51.0%; Score 74; DB 1; Length 222;  
 CC Best Local Similarity 42.9%; Pred. No. 0.00058;  
 CC Matches 12; Conservative 10; Mismatches 6; Indels 0; Gaps 0;  
 CC  
 CC QY 1 KYSLPEIDYEFSAPEYISQINIEIXYT 28  
 CC Db 25 KYSLPDLPIYDYGALPHINQIQLHHS 52  
 CC  
 CC RESULT 18  
 CC SODM\_MOUSE STANDARD; PRT; 222 AA.  
 CC ID SODM\_MOUSE  
 CC AC P09671; 064670;  
 CC DT 01-MAR-1989 (Rel. 10, Created)  
 CC DT 15-JUL-1998 (Rel. 36, Last sequence update)

[illegible]

ID	SODM_HUMAN	STANDARD:	PRT:	222 AA.
AC	P04179;	Q16792;	Q9P223;	
DT	20-MAR-1987 (Rel. 04, Created)			
DT	01-APR-1990 (Rel. 14, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Superoxide dismutase [Mn], mitochondrial precursor (EC 1.15.1.1).			
GN	SOD2.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxId=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=89076921; PubMed=2462451;			
RA	Wajape J.R., Clark J.C., Burnhans M.S., Kropp K.E., Korfhagen T.R.,			
RA	Whiteett J.A.;			
RT	"Synthesis and processing of the precursor for human			
RT	mangano-superoxide dismutase."			
RL	Biochim. Biophys. Acta 994:30-36(1989).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88067716; PubMed=3684581;			
RA	Beck Y., Oren R., Amit B., Levanon A., Gorecki M., Hartman J.R.;			
RA	"Human Mn superoxide dismutase cDNA sequence."			
RT	Nucleic Acids Res. 15:9076-9076(1987).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88289364; PubMed=3399391;			
RA	Heckl K.;			
RT	"Isolation of cDNAs encoding human manganese superoxide dismutase."			
RL	Nucleic Acids Res. 16:6224-6224(1988).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Liver;			
RX	MEDLINE=88152250; PubMed=2831093;			
RA	Ho Y.-S., Crapo J.D.;			
RT	"Isolation and characterization of complementary DNAs encoding human			
RT	manganese-containing superoxide dismutase."			
RL	FEBS Lett. 229:256-260(1988).			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Colon;			
RX	MEDLINE=91105727; PubMed=1988135;			
RA	St Clair D.K., Holland J.C.;			
RT	"Complementary DNA encoding human colon cancer manganese superoxide			
RT	dismutase and the expression of its gene in human cells."			
RL	Cancer Res. 51:939-943(1991).			
RN	[6]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=91027939; PubMed=1699607;			
RA	Church S.L.;			
RT	"Manganese superoxide dismutase: nucleotide and deduced amino acid			
RT	sequence of a cDNA encoding a new human transcript."			
RL	Biochim. Biophys. Acta 1087:250-252(1990).			
RN	[7]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=95217333; PubMed=7702755;			
RA	Wan X.S., Devalaraja K.N., St Clair D.K.;			
RT	"Molecular structure and organization of the human manganese			
RT	superoxide dismutase gene."			
RL	DNA Cell Biol. 13:1127-1136(1994).			
RN	[8]			
RP	SEQUENCE OF 25-222.			
RX	MEDLINE=85030346; PubMed=6386798;			
RA	Barra D., Schinella M.E., Sirmaco M., Bannister J.V., Bannister W.H.,			
RA	Rottlio G., Bosca F.;			
RT	"The primary structure of human liver manganese superoxide			
RT	dismutase."			
RL	J. Biol. Chem. 259:12595-12601(1984).			
RN	[9]			

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sb-sib.ch](mailto:license@sb-sib.ch)).

-----

CC EMBL; L28808; AAA31401.1; -.  
DR HSSP; P04179; IAP6.  
DR InterPro; IPR001189; SODismutase.  
DR Pfam; PF00081; sodfe; 1.  
DR Pfam; PF02777; sodfe.C; 1.  
DR ProDom; PD000475; SODismutase; 1.  
DR PROSITE; PS00088; SOD\_MN; 1.  
KW Oxidoreductase; Manganese; Mitochondrion; Transil peptide.  
FT NON TER 1 1  
FT TRANSIT <1 5 MITOCHONDRION (BY SIMILARITY).  
FT CHAIN 6 >202 SUPEROXIDE DISMUTASE [MN].  
FT METAL 31 31 MANGANESE (BY SIMILARITY).  
FT METAL 79 79 MANGANESE (BY SIMILARITY).  
FT METAL 164 164 MANGANESE (BY SIMILARITY).  
FT METAL 168 168 MANGANESE (BY SIMILARITY).  
FT NON TER 202 202  
SEQUENCE 202 AA; 22656 MW; 6F1BB8DA15C33AA7 CRC64;

Query Match 53.1%; Score 77; DB 1; Length 202;  
Best Local Similarity 46.4%; Pred. No. 0.00018;  
Matches 13; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

Qy 1 KYSLPELYERSATEPYISGOINEIXYT 28  
Db 6 KHSLPDLPYDYGALPHINQIMELHHS 33

RESULT 14  
SODF\_METU STANDARD; PRT; 202 AA.  
ID SODF\_METU  
AC P23744;  
DT 01-NOV-1991 (Rel. 20, Created)  
DT 01-NOV-1991 (Rel. 20, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Superoxide dismutase [Mn/Fe] (EC 1.15.1.1).  
GN SODB.  
OS Methylomonas J.  
OC Bacteria; Proteobacteria; gamma subdivision; Methylococcaceae;  
OC Methylomonas.  
OX NCBI\_TaxID=32038;  
CC [1]  
CC SEQUENCE.  
RP MEDLINE=9182736; PubMed=1848999;  
RA Matsumoto T., Terauchi K., Isobe T., Matsunaka K., Yamakura F.;  
"Iron- and manganese-containing superoxide dismutases from  
Methylomonas J: identity of the protein moiety and amino acid  
sequence";  
RL Biochemistry 30:3210-3216(1991).  
RN [2]  
RN CHARACTERIZATION, AND SEQUENCE OF 1-32.  
RX MEDLINE=91301511; PubMed=1906419;  
RA Yamakura F., Matsumoto T., Terauchi K.;  
RT "Isolation of Mn-SOD and low active Fe-SOD from Methylomonas J;  
consisting of identical proteins";  
RL Free Radic. Res. Commun. 12:329-334(1991).  
CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE  
CC CELLS AND WHICH ARE TOXIC TO BIOLOGICAL SYSTEMS.  
CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).  
CC -1- COFACTOR: ACTIVE WITH EITHER MANGANESE OR IRON AS A COFACTOR.  
CC BOTH TYPES OF SODS ARE COMPOSED OF A SINGLE APOPROTEIN SYNTHESIZED  
CC IN CELLS GROWN IN EITHER THE FE-RICH OR THE MN-RICH MEDIUM. THE  
CC FE-SOD HAS A LOWER ACTIVITY THAN THE MN-SOD.  
CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE  
CC FAMILY.  
DR PIR; A38461; A38461.  
DR HSSP; P00448; IYEW.  
DR InterPro; IPR001189; SODismutase.  
DR Pfam; PF00081; sodfe; 1.

DR Pfam; PF02777; sodfe.C; 1.  
DR ProDom; PD000475; SODismutase; 1.  
DR PROSITE; PS00088; SOD\_MN; 1.  
KW Oxidoreductase; Manganese; Iron.  
FT METAL 26 26 MANGANESE OR IRON (BY SIMILARITY).  
FT METAL 80 80 MANGANESE OR IRON (BY SIMILARITY).  
FT METAL 163 163 MANGANESE OR IRON (BY SIMILARITY).  
FT METAL 167 167 MANGANESE OR IRON (BY SIMILARITY).  
FT METAL 167 167 MANGANESE OR IRON (BY SIMILARITY).  
SEQUENCE 202 AA; 22363 MW; 321A8B4697A5B58 CRC64;

Query Match 52.4%; Score 76; DB 1; Length 202;  
Best Local Similarity 51.9%; Pred. No. 0.00026;  
Matches 14; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

Qy 2 YSLPELYERSATEPYISGOINEIXYT 28  
Db 2 YTLPELDYATLPHIDQTMELHHT 28

RESULT 15  
SODF\_ACIAM STANDARD; PRT; 211 AA.  
ID SODF\_ACIAM  
AC Q9P9J3;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Superoxide dismutase [Fe] (EC 1.15.1.1).  
GN SOD.  
OS Acididans ambivalens (Desulfurolobus ambivalens).  
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;  
OC Acididans.  
OX NCBI\_TaxID=2283;  
CC [1]  
CC SEQUENCE FROM N.A.  
RX MEDLINE=21026956; PubMed=11154067;  
RA Kardinahl S., Anemuller S., Schaefer G.;  
RT "The hyper-thermostable Fe-superoxide dismutase from the Archaeon  
RT Acididans ambivalens: characterization, recombinant expression,  
RT crystallization and effects of metal exchange";  
RL Biol. Chem. 381:1089-1101(2000).  
CC -1- FUNCTION: Destroys radicals which are normally produced within the  
CC cells and which are toxic to biological systems (By similarity).  
CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).  
CC -1- COFACTOR: Iron (By similarity).  
CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE  
CC FAMILY.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sb-sib.ch](mailto:license@sb-sib.ch)).

-----

CC EMBL; AF236110; AAF36989.1; -.  
DR HSSP; P80857; 1SSS.  
DR InterPro; IPR001189; SODismutase.  
DR Pfam; PF00081; sodfe; 1.  
DR Pfam; PF02777; sodfe.C; 1.  
DR ProDom; PD000475; SODismutase; 1.  
DR PROSITE; PS00088; SOD\_MN; FALSE\_NEG.  
KW Oxidoreductase; Iron.  
FT METAL 34 34 IRON (BY SIMILARITY).  
FT METAL 82 82 IRON (BY SIMILARITY).  
FT METAL 171 171 IRON (BY SIMILARITY).  
FT METAL 175 175 IRON (BY SIMILARITY).  
SEQUENCE 211 AA; 24342 MW; 5A88PFA400F77065 CRC64;

Query Match 51.0%; Score 74; DB 1; Length 211;  
Best Local Similarity 51.9%; Pred. No. 0.00055;  
Matches 14; Conservative 3; Mismatches 10; Indels 0; Gaps 0;







OX NCBI\_TaxID=1766;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 6841;  
RA MEDLINE=96102872; PubMed=8586279;  
RT Menendez M.C., Domenech P., Prieto J., Garcia M.J.;  
RT "Cloning and expression of the Mycobacterium fortuitum superoxide  
dismutase gene";  
RL FEMS Microbiol. Lett. 134:273-278(1995).  
CC -1- FUNCTION: Destroys radicals which are normally produced within the  
CC cells and which are toxic to biological systems.  
CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).  
CC -1- COFACTOR: Manganese.  
CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE  
CC FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X70914; CAA50266.1; -.  
CC HSSP; PL7670; 1IDS.  
CC InterPro; IPR00189; SODismutase.  
CC Pfam; PF00081; sodfe; 1.  
CC DR Pfam; PF02777; sodfe\_C; 1.  
CC DR Prodom; PD000475; SODismutase; 1.  
CC DR PROSITE; PS00088; SOD\_MN; 1.  
CC KM Oxidoreductase; Manganese.  
CC FT INIT MET 0 BY SIMILARITY.  
CC FT METAL 27 27 MANGANESE (BY SIMILARITY).  
CC FT METAL 75 75 MANGANESE (BY SIMILARITY).  
CC FT METAL 159 159 MANGANESE (BY SIMILARITY).  
CC FT METAL 163 163 MANGANESE (BY SIMILARITY).  
CC SQ SEQUENCE 206 AA; 22833 MW; 500625A8B9321246 CRC64;  
Query Match 68.3%; Score 99; DB 1; Length 206;  
Best Local Similarity 57.1%; Pred. No. 8.9e-08;  
Matches 16; Conservative 9; Mismatches 3; Indels 0; Gaps 0;  
QY 1 KYSLPELDYFSPATEPTISQGINEIXYT 28  
Db 2 EYTLPLDIDYGALEPHISQGINELHHS 29  
ULI 7  
M\_MYCSM STANDARD; PRT; 206 AA.  
AC P53643; Q98612;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Superoxide dismutase [Mn] (EC 1.15.1.1).  
OS SODA OR SOD  
OS Mycobacterium smegmatis.  
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;  
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1772;  
RN [1]  
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.  
RC STRAIN=1-2C;  
RA MEDLINE=99134360; PubMed=9933629;  
RA Harth G., Horwitz M.A.;  
RT "Export of recombinant Mycobacterium tuberculosis superoxide dismutase  
RT is dependent upon both information in the protein and mycobacterial  
RT export machinery. A model for studying export of leaderless proteins  
RT by pathogenic mycobacteria";  
RL J. Biol. Chem. 274:4281-4292(1999).  
RN [2]  
RP REVISIONS TO 116 AND 201.

RA Tullius M.V., Harth G., Horwitz M.A.;  
RN Submitted (SEP-2001) to the EMBL/Genbank/DBJ databases.  
RL [3]  
RP SEQUENCE OF 27-164 FROM N.A.  
RC STRAIN=NCTC 10265 / ATCC 14468 / W-113;  
RA Bull T.J., Shanson D.C., Archard L.C.;  
RT "Rapid identification of mycobacteria from AIDS patients by capillary  
RT electrophoretic profiling of amplified SOD gene";  
RL J. Clin. Pathol. Clin. Mol. Pathol. 48:124-132(1995).  
CC -1- FUNCTION: Destroys radicals which are normally produced within the  
CC cells and which are toxic to biological systems.  
CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).  
CC -1- COFACTOR: Manganese.  
CC -1- SUBUNIT: HOMOTETRAMER.  
CC -1- SUBCELLULAR LOCATION: PARTIALLY SECRETED.  
CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE  
CC FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AF061031; AAD15825.2; -.  
CC HSSP; Z48214; CAA8247.1; -.  
CC DR HSSP; PL7670; 1IDS.  
CC DR InterPro; IPR00189; SODismutase.  
CC DR Pfam; PF00081; sodfe; 1.  
CC DR Pfam; PF02777; sodfe\_C; 1.  
CC DR Prodom; PD000475; SODismutase; 1.  
CC DR PROSITE; PS00088; SOD\_MN; 1.  
CC KM Oxidoreductase; Manganese.  
CC FT INIT MET 0 BY SIMILARITY.  
CC FT METAL 27 27 MANGANESE (BY SIMILARITY).  
CC FT METAL 75 75 MANGANESE (BY SIMILARITY).  
CC FT METAL 159 159 MANGANESE (BY SIMILARITY).  
CC FT METAL 163 163 MANGANESE (BY SIMILARITY).  
CC SQ SEQUENCE 206 AA; 22804 MW; B97A1AD0374AC9 CRC64;  
Query Match 68.3%; Score 99; DB 1; Length 206;  
Best Local Similarity 57.1%; Pred. No. 8.9e-08;  
Matches 16; Conservative 9; Mismatches 3; Indels 0; Gaps 0;  
QY 1 KYSLPELDYFSPATEPTISQGINEIXYT 28  
Db 2 EYTLPLDIDYGALEPHISQGINELHHS 29  
RESULT 8  
SODM\_MYCLB STANDARD; PRT; 206 AA.  
AC P13367;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Superoxide dismutase [Mn] (EC 1.15.1.1).  
OS SODA OR ML0072.  
OS Mycobacterium leprae.  
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;  
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1769;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=90045970; PubMed=2682526;  
RA Thangaraj H.S., Lamb F.I., Davis E.O., Colston M.J.;  
RT "Nucleotide and deduced amino acid sequence of Mycobacterium leprae  
RT manganese superoxide dismutase";  
RL Nucleic Acids Res. 17:8378-8378(1989).  
RN [2]  
RP SEQUENCE FROM N.A.



```

CC      cells and which are toxic to biological systems.
CC      -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC      -1- COFACTOR: Manganese (By similarity).
CC      -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE
CC      FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; D13288; BAA28850.1; -.
CC      HSSP; P17670; 1IDS.
CC      InterPro; IPR001189; SODismutase.
CC      Pfam; PF000681; sodfe, 1.
CC      Pfam; PF02777; sodfe_C, 1.
CC      ProDom; PD000475; SODismutase; 1.
CC      PROSITE; PS00088; SOD_MN; 1.
CC      OXidoreductase; Manganese.
CC      KM
CC      INIT MET 0 0 BY SIMILARITY.
CC      FT METAL 27 27 MANGANESE (BY SIMILARITY).
CC      FT METAL 75 75 MANGANESE (BY SIMILARITY).
CC      FT METAL 159 159 MANGANESE (BY SIMILARITY).
CC      FT METAL 163 163 MANGANESE (BY SIMILARITY).
CC      SQ SEQUENCE 206 AA; 23031 MW; 92A063471FA9F22F CRC64;

Query Match 69.0%; Score 100; DB 1; Length 206;
Best Local Similarity 60.7%; Pred. No. 6.3e-08;
Matches 17; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

Qy 1 KYSLPELDYEFSTPEYISGQINEIXYT 28
Db 2 EYTLPDMDYEALEPHISGQINEIHT 29

RESULT 5
SODM_YEAST STANDARD; PRT; 233 AA.
ID SODM_YEAST
AC P00447;
DT 21-UTL-1986 (Rel. 01, Created)
DT 23-OCT-1986 (Rel. 02, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Superoxide dismutase [Mn], mitochondrial precursor (EC 1.15.1.1).
DE SOD2 OR YHR008C.
GN Saccharomyces cerevisiae (Baker's yeast).
GN Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
GN Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
GN NCBI_TaxId=4932;
GN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=85127011; Pubmed=3882422;
RX Warren C.A.M., van Loon A.P.G.M., Oudehoorn P., van Steeg H.,
RA Grivell L.A., Slater E.C.;
RA "Nucleotide sequence analysis of the nuclear gene coding for
RA manganese superoxide dismutase of yeast mitochondria, a gene
RA previously assumed to code for the Rieske iron-sulphur protein.";
RA Eur. J. Biochem. 147:153-161(1985).
RN [2]
RN SEQUENCE FROM N.A.
RP STRAIN=S288c / AB972;
RX MEDLINE=94378003; Pubmed=8091229;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
RA Du Z., Favello A., Fulton L., Gattung S., Geisel C., Kirksen J.,
RA Kucada T., Hillier L., Jier M., Johnston L., Langston Y.,
RA Latreille P., Louis E.J., Macri C., Mardis E., Menezes S., Mouser L.,
RA Nham M., Rifkin L., Riles L., St Peter H., Trevasaki E., Vaughan K.,
RA Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,
RA Vaudin M.;
RA "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
RA VIII.";

```

```

RL Science 265:2077-2082(1994).
RN [3]
RN SEQUENCE OF 1-39 FROM N.A.
RX MEDLINE=89211942; Pubmed=3072251;
RA Schrank I.S., Sims P.F., Oliver S.G.;
RT "Functional expression of the yeast Mn-superoxide dismutase gene in
RT Escherichia coli requires deletion of the signal peptide sequence.";
RL Gene 73:121-130(1988).
RN [4]
RN SEQUENCE OF 27-233.
RA Dittlow C., Johansen J.T., Martin B.M., Svendsen I.;
RT "The complete amino acid sequence of manganese-superoxide dismutase
RT from Saccharomyces cerevisiae.";
RL Carlsberg Res. Commun. 47:81-91(1982).
CC -1- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems.
CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -1- COFACTOR: Manganese.
CC -1- SUBUNIT: Homotetramer.
CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE
CC FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; X02156; CAA26092.1; -.
CC      EMBL; U10400; AAB68939.1; -.
CC      EMBL; M24079; AAB35065.1; -.
CC      DR PIR; A00521; DSBVN.
CC      DR PIR; S46785; S46785.
CC      DR HSSP; P04179; 1ABM.
CC      DR SGD; S0001050; SOD2.
CC      DR InterPro; IPR001189; SODismutase.
CC      DR Pfam; PF00081; sodfe, 1.
CC      DR Pfam; PF02777; sodfe_C, 1.
CC      DR ProDom; PD000475; SODismutase; 1.
CC      PROSITE; PS00088; SOD_MN; 1.
CC      OXidoreductase; Manganese; Mitochondrion; Transic peptide.
CC      KM
CC      TRANSIT 1 26 MITOCHONDRION.
CC      FT CHAIN 27 233 SUPEROXIDE DISMUTASE [MN].
CC      FT METAL 52 52 MANGANESE (BY SIMILARITY).
CC      FT METAL 107 107 MANGANESE (BY SIMILARITY).
CC      FT METAL 194 194 MANGANESE (BY SIMILARITY).
CC      FT METAL 198 198 MANGANESE (BY SIMILARITY).
CC      SQ SEQUENCE 233 AA; 25774 MW; 88A9391FBB31D06E CRC64;

Query Match 69.0%; Score 100; DB 1; Length 233;
Best Local Similarity 64.3%; Pred. No. 7.2e-08;
Matches 18; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 1 KYSLPELDYEFSTPEYISGQINEIXYT 28
Db 27 KYTLPDMDYEALEPHISGQINEIHT 54

RESULT 6
SODM_MYCFO STANDARD; PRT; 206 AA.
ID SODM_MYCFO
AC O59519;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Superoxide dismutase [Mn] (EC 1.15.1.1).
DE SOD4 OR SOD.
GN Mycobacterium fortuitum.
OS Bacteria; Actinobacteria; Actinobacteriales; Actinobacteridae;
OC Actinomycetales; Corynebacteriaceae; Mycobacterium.

```



Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KYSLPELDYERSATEPYISGOINEIXYT 28  
 |||||  
 DB 35 KYSLPELDYERSATEPYISGOINEIHYT 62

RESULT 2  
 SODM\_NOCAS STANDARD; PRT; 206 AA.  
 AC P53651;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Superoxide dismutase [Mn] (EC 1.15.1.1).  
 GN SODA OR SOD.  
 OS Nocardiasteroides.  
 OS Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;  
 OC Actinomycetales; Corynebacterineae; Nocardiaceae; Nocardia.  
 OX NCBI\_TaxID=1824;  
 [1]  
 SEQUENCE FROM N.A.  
 AC STRAIN=GUH2;  
 RX MEDLINE=96060854; PubMed=7590304;  
 RA Alcedor D.J., Chapman G.D., Beaman B.D.;  
 RT "Isolation, sequencing and expression of the superoxide dismutase-  
 encoding gene (sod) of Nocardia asteroides strain GUH-2.";  
 RL Gene 164:143-147(1995).  
 CC -1- FUNCTION: Destroys radicals which are normally produced within the  
 cells and which are toxic to biological systems.  
 CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).  
 CC -1- COFACTOR: Manganese (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE  
 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; U02341; AAA91964.1; -.  
 DR HSSP; P17670; IIDS.  
 DR InterPro; IPR001189; SODismutase.  
 DR Pfam; PF00081; sodfe; 1.  
 DR ProDom; PD000475; SODismutase; 1.  
 PROSITE; PS00088; SOD\_MN; 1.  
 OXidoreductase; Manganese.  
 FT INIT MET 0  
 FT METAL 27 27 BY SIMILARITY.  
 FT METAL 75 75 MANGANESE (BY SIMILARITY).  
 FT METAL 159 159 MANGANESE (BY SIMILARITY).  
 FT METAL 163 163 MANGANESE (BY SIMILARITY).  
 SQ SEQUENCE 206 AA; 22823 MW; FA2BCF27EB22CBDCD CRC64;

Query Match 71.0%; Score 103; DB 1; Length 206;  
 Best Local Similarity 60.7%; Pred. No. 2.2e-08;  
 Matches 17; Conservative 9; Mismatches 2; Indels 0; Gaps 0;

QY 1 KYSLPELDYERSATEPYISGOINEIXYT 28  
 :||:|||||:|||||:|  
 DB 2 EYLPDLDYDYSALPHISGOINEIHYT 29

RESULT 3  
 SODM\_MYCAV STANDARD; PRT; 206 AA.  
 AC P47201;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Superoxide dismutase [Mn] (EC 1.15.1.1).  
 GN SODA OR SOD.  
 OS Mycobacterium avium.  
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;  
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1764;  
 [1]  
 SEQUENCE FROM N.A.

RP STRAIN=724;  
 RX MEDLINE=96276149; PubMed=8692009;  
 RA Escuyer V.E., Haddad N., Frenel C., Berche P.;  
 RT "Molecular characterization of a surface-exposed superoxide dismutase  
 of Mycobacterium avium.";  
 RL Microb. Pathog. 20:41-55(1996).  
 CC -1- FUNCTION: Destroys radicals which are normally produced within the  
 cells and which are toxic to biological systems.  
 CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).  
 CC -1- COFACTOR: Manganese.  
 CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE  
 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; U11550; AAB08770.1; -.  
 DR HSSP; P17670; IIDS.  
 DR InterPro; IPR001189; SODismutase.  
 DR Pfam; PF00081; sodfe; 1.  
 DR ProDom; PD000475; SODismutase; 1.  
 DR PROSITE; PS00088; SOD\_MN; 1.  
 KW Oxidoreductase; Manganese.  
 FT INIT MET 0  
 FT METAL 27 27 BY SIMILARITY.  
 FT METAL 75 75 MANGANESE (BY SIMILARITY).  
 FT METAL 159 159 MANGANESE (BY SIMILARITY).  
 FT METAL 163 163 MANGANESE (BY SIMILARITY).  
 SQ SEQUENCE 206 AA; 22912 MW; D6B49083AEBF3C98 CRC64;

Query Match 69.7%; Score 101; DB 1; Length 206;  
 Best Local Similarity 60.7%; Pred. No. 4.4e-08;  
 Matches 17; Conservative 9; Mismatches 2; Indels 0; Gaps 0;

QY 1 KYSLPELDYERSATEPYISGOINEIXYT 28  
 :||:|||||:|||||:|  
 DB 2 EYLPDLDYDYSALPHISGOINEIHYT 29

RESULT 4  
 SODM\_MYCLP STANDARD; PRT; 206 AA.  
 AC O86165;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Superoxide dismutase [Mn] (EC 1.15.1.1).  
 GN SODA OR SOD.  
 OS Mycobacterium lepreum.  
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;  
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=64667;  
 [1]  
 SEQUENCE FROM N.A.  
 RC STRAIN=Hawaiian.  
 RA Nakamura M.;  
 RL Submitted (SEP-1992) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: Destroys radicals which are normally produced within the

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 9, 2003, 13:38:54; Search time 3.85321 Seconds

(without alignments)  
322.923 Million cell updates/sec

Title: US-09-987-190-2

Perfect score: 145

Sequence: 1 KYSLPELDYFSPATEPYISQINEIXYTX 30

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Database: SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Length	ID	Description
1	141	97.2	234 1	SODM_CANAL
2	103	71.0	206 1	SODM_NOCAS
3	101	69.7	206 1	SODM_MYCAV
4	100	69.0	206 1	SODM_MYCAB
5	100	69.0	233 1	SODM_YEAST
6	99	68.3	206 1	SODM_MYCFO
7	99	68.3	206 1	SODM_MYCSM
8	98	67.6	206 1	SODM_MYCLE
9	94	64.8	207 1	SODM_MYCTU
10	91	62.8	201 1	SODM_PROFR
11	83	57.2	245 1	SODM_NEUCR
12	80	55.2	222 1	SODM_HORSE
13	77	52.4	202 1	SODM_RABIT
14	76	52.4	202 1	SODF_METU
15	74	51.0	211 1	SODF_ACIAJ
16	74	51.0	222 1	SODM_BOVIN
17	74	51.0	222 1	SODM_HUMAN
18	74	51.0	222 1	SODM_MOUSE
19	74	51.0	222 1	SODM_RAT
20	73	50.3	210 1	SODF_SULSO
21	72	49.7	211 1	SODF_PYRAE
22	72	49.0	228 1	SODM_CAPAN
23	71	49.0	210 1	SODM_ASFPU
24	71	48.3	233 1	SODM_HEYBR
25	70	48.3	203 1	SODM_BORBU
26	70	48.3	207 1	SODM_CHLPN
27	70	48.3	231 1	SODM_ARATH
28	70	48.3	233 1	SODM_PEA
29	69	47.6	281 1	SODF_BACSU
30	69	47.6	238 1	SODM_NICPL
31	68	46.9	200 1	SODM_AGABT
32	68	46.9	210 1	SODF_SULAC
33	68	46.9	211 1	SODM_CAVPO

34	68	46.9	212 1	SODF_STRCO	OS1917 streptomyc
35	68	46.9	232 1	SODM_MAIZE	P41978 zea mays (m
36	68	46.9	233 1	SODO_MAIZE	P41979 zea mays (m
37	67	46.2	23 1	SODM_RANCA	P36215 rana catesb
38	67	46.2	203 1	SODM_RANCP	P36554 xanthomonas
39	67	46.2	214 1	SODM_HAEIN	P43725 haemophilus
40	67	46.2	224 1	SODM_CHAFE	O96347 charybdis f
41	67	46.2	233 1	SODF_MAIZE	P41980 zea mays (m
42	67	46.2	235 1	SODM_MAIZE	P09223 zea mays (m
43	66	45.5	207 1	SODM_YEREN	P53655 yersinia en
44	65	44.8	200 1	SODM_YEREN	O92429 gandodema m
45	65	44.8	202 1	SODF_METTM	Q60036 methanobact

## ALIGNMENTS

RESULT 1  
ID SODM\_CANAL STANDARD; PRT; 234 AA.  
AC 013401;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Superoxide dismutase [Mn], mitochondrial precursor (EC 1.15.1.1).  
GN SOD2.  
OS Candida albicans (Yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.  
OX NCBI\_TaxID=5476;  
RN [1]  
RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND CHARACTERIZATION.  
RX MEDLINE=9177423; PubMed=10076057;  
RA Rile G.E., Hwang C.S., Brady M.J., Kim S.T., Kim Y.R., Hub W.K.,  
RA Baek Y.U., Lee B.H., Lee J.S., Kang S.O.,  
RT "Manganese-containing superoxide dismutase and its gene from Candida  
albicans.";  
RL Biochim. Biophys. Acta 1426:409-419 (1999).  
CC -!- FUNCTION: Destroys radicals which are normally produced within the  
cells and which are toxic to biological systems.  
CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).  
CC -!- COFACTOR: Manganese.  
CC -!- SUBUNIT: HOMOTETRAMER.  
CC -!- SUBCELLULAR LOCATION: Mitochondrial matrix.  
CC -!- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE  
FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL: AF031478; AAB86583.1; -.  
DR HSSP: P04179; IABM.  
DR InterPro: IPR001189; SODismutase.  
DR Pfam: PF00081; sodefe, 1.  
DR Pfam: PF02777; sodefe\_C, 1.  
DR ProDom: PD000475; SODismutase, 1.  
DR PROSITE: PS00088; SOD\_MN, 1.  
KW Oxidoreductase; Manganese; Mitochondrion; Transit peptide.  
FT TRANSIT 1 34  
FT CHAIN 35 234  
FT METAL 60 60  
FT METAL 108 108  
FT METAL 198 198  
FT METAL 202 202  
SQ SEQUENCE 234 AA; 26173 MW; EBFPC2D765CID9C1 CR664;  
Query Match 97.2%; Score 141; DB 1; Length 234;  
Best Local Similarity 96.4%; Pred. No. 4.5e-14;

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 8, 2003, 02:09:43 ; Search time 278.532 Seconds

(without alignments)  
1744.376 Million cell updates/sec

Title: US-09-987-190-2

Perfect score: 145

Sequence: 1 KYSLELDYFSAATEPYISGQINIXYTX 30

Scoring table:

BLOSUM62	Xgapop 10.0	Xgapext 0.5
	Ygapop 10.0	Ygapext 0.5
	Dgapop 6.0	Dgapext 7.0
	Delop 6.0	Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

```
-MODEL=frame.p2n.model -DEV=x1h
-O=/cgn2.1/USPTO.spool/US09987190/runat.02042003.092633.19234/app.query.fasta.1.526
-DB=EST -QFMT=fastap -SUFFIX=1st -MINMATCH=0.1 -DOOPCL=0 -DDEPXT=0
-UNIT5=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=DCT -THR MAX=100 -THR MIN=0 -ALIGN=40 -MODE=LOCAL
-OUTFMT=ptc -NOM=ext -HEAPSIZE=500 -MILLEN=0 -MAXLEN=200000000
-USER=US09987190 @CNC 1.1 2013 @runat.02042003.092633.19234 -NCPU=6 -ICPU=3
-NO_XIPXY -NO_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-MARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database :

```
EST:*
1: em_estba:*
2: em_esthum:*
3: em_estcin:*
4: em_estmu:*
5: em_estcov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_vit:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	141	97.2	862	17	CNS07DTU
2	92	63.4	519	17	A2929307
3	90	62.1	464	14	B0087025
4	83	57.2	231	9	A1319465
5	83	57.2	265	9	A1320309
6	83	57.2	265	9	AW717655
7	83	57.2	267	9	A1318805
8	83	57.2	267	10	AW723904
9	83	57.2	269	9	A1321638
10	83	57.2	336	9	A1320505
11	83	57.2	339	9	A1318710
12	83	57.2	339	9	A1319577
13	83	57.2	340	10	AW715499
14	83	57.2	343	10	AW724497
15	83	57.2	401	10	AW724781
16	83	57.2	429	10	AW715120
17	83	57.2	429	10	AW715952
18	82	56.6	592	13	B1321452
19	82	56.6	689	13	BM266232
20	80	55.2	466	13	BM367690
21	79	54.5	306	10	AW733168
22	79	54.5	419	13	B1498255
23	79	54.5	420	10	AW756619
24	79	54.5	425	10	BE210200
25	79	54.5	448	9	A1460811
26	79	54.5	448	9	A1460823
27	79	54.5	470	10	AM099772
28	79	54.5	463	10	AM099772
29	79	54.5	521	12	BE802643
30	79	54.5	552	10	BE211438
31	79	54.5	564	13	B1973216
32	79	54.5	571	13	BM521591
33	79	54.5	578	14	B0094786
34	79	54.5	592	10	BE210646
35	79	54.5	592	13	BM521797
36	79	54.5	597	12	BF067640
37	79	54.5	598	10	AM234447
38	79	54.5	670	10	BE330334
39	79	54.5	675	13	B1263359
40	79	54.5	683	13	B1942249
41	79	54.5	706	10	AM184867
42	78	53.8	423	10	AM099781
43	78	53.8	447	10	BE122013
44	78	53.8	476	12	BE726243
45	78	53.8	599	13	B1531113

#### ALIGNMENTS

RESULT 1  
CNS07DTU 862 bp DNA linear GSS 08-JUL-2001  
LOCUS T7 end of clone BD0AA015F01 of library BD0AA from strain CBS 94 of  
DEFINITION Candida tropicalis, genomic survey sequence.  
ACCESSION AL440909.1 GI:12224320  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Candida tropicalis.  
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
Saccharomycetales; mitosporic Saccharomycetales; Candida.  
REFERENCE  
1 (bases 1 to 862)  
Souleir,J.L., Aigle,M., Artiguenave,F., Blandin,G.,  
Bollotin-Fukuhara,M., Bon,E., Brotlier,P., Casaregola,S.,

de-Montigny, J., Dujon, B., Durrens, P., Lepingle, A., Llorente, B.,  
Maurer, A., Neugebäude, C., Ozier-Kalogeropoulos, O., Potier, S.,  
Saitou, M., Tekai, F., Toffano-Nioche, C., Wesolowski-Louvel, M.,  
Wincker, P. and Weissbach, J.  
Genomic exploration of the hemiascomycetous yeasts: 1. A set of  
yeast species for molecular evolution studies  
FEMS Lett. 487 (1), 3-12 (2000)

JOURNAL MEDLINE  
20584711  
1152876  
PUBMED

REFERENCE  
2 (bases 1 to 862)  
Blandin, G., Ozier-Kalogeropoulos, O., Wincker, P., Artiguenave, F. and  
Dujon, B.  
Genomic exploration of the hemiascomycetous yeasts: 16. Candida  
tropicalis  
FEMS Lett. 487 (1), 91-94 (2000)

JOURNAL MEDLINE  
20584726  
1152891  
PUBMED

REFERENCE  
3 (bases 1 to 862)  
Genoscope.  
Direct Submission  
Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,  
2 rue Gaston Cremieux, CP 5706, 91057 Evry cedex, FRANCE. (E-mail :  
sequef@genoscope.cns.fr - Web :  
http://www.genoscope.cns.fr)  
This GSS is part of a random genomic sequencing program of thirteen  
yeast species: *Saccharomyces bayanus* var. *uvarum*, *Saccharomyces  
exiguus*, *Saccharomyces servazzii*, *Zygosaccharomyces rouxii*,  
*Saccharomyces kluyveri*, *Kluyveromyces thermocolarans*, *Kluyveromyces  
lactis* var. *lactis*, *Kluyveromyces marxianus* var. *marxianus*, *Pichia  
angusta*, *Debaryomyces hansenii* var. *hansenii*, *Pichia sorbitophila*,  
*Candida tropicalis* and *Yarrowia lipolytica*. Genomic inserts of 3 to  
5 kb were prepared and both extremities were sequenced. See  
keywords for description of this sequence and for the sequence of  
the other extremity of this insert.

## FEATURES

source

misc\_feature

1..862  
/organism="Candida tropicalis"  
/strain="CBS 94"  
/db\_xref="taxon:5482"  
/clone\_1b="BD0A015F01"  
/clone\_1lb="BD0A0A"  
/note="end : T7"  
<124..>726  
/note="similar to *Saccharomyces cerevisiae* ORF YHR008C [SOD2 : superoxide dismutase (Mn) precursor, mitochondrial]"  
/evidence="not experimental"  
BASE COUNT 278 a 117 c 140 g 322 t 5 others  
ORIGIN

## Alignment Scores:

4.5e-13 Length: 862  
Score: 141.00 Matches: 27  
Percent Similarity: 96.43% Conservative: 0  
Best Local Similarity: 96.43% Mismatches: 1  
Query Match: 97.24% Indels: 0  
DB: 17 Gaps: 0

US-09-987-190-2 (1-30) x CNS07DTJ (1-862)

Qy 1 LyeTySerIeuProGluLeuAapTyGluPheserAlaThrGluProTyrllesergly 20  
Db 130 AATATATAGTTTACAGAAATTAAGATTATTTCTCTACTGAACATATATTTTCGA 189  
Qy 21 Glnlleasngluile\*\*\*TyrThr 28  
Db 190 CAAATTAATGAAATTCATTAATCT 213

RESULT 2  
AZ929307 519 bp DNA linear GSS 01-APR-2001  
LOCUS  
DEFINITION 479.dif24g04.g1 Saccharomyces kluyveri Saccharomyces kluyveri  
ACCESSION genomic clone 479.dif24g04.g1, DNA sequence.  
AZ929307

VERSION  
KEYWORDS  
SOURCE  
ORGANISM

## REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Department of Genetics  
Washington University Medical School  
Box 8232, 4566 Scott Ave., St. Louis, MO 63110, USA  
Tel: 314 362 2735  
Fax: 314 362 7855  
Email: mj@genetics.wustl.edu  
Class: random plasmid subclone.  
Location/Qualifiers

## FEATURES

source

1..519  
/organism="Saccharomyces kluyveri"  
/strain="NRRL Y-12651 (CBS 3082)"  
/db\_xref="taxon:4934"  
/clone="479.dif24g04.g1"  
/clone\_1lb="Saccharomyces kluyveri"  
/note="Random genomic sequence"  
BASE COUNT 147 a 143 c 83 g 146 t

## Alignment Scores:

3.15e-05 Length: 519  
Score: 92.00 Matches: 16  
Percent Similarity: 82.14% Conservative: 7  
Best Local Similarity: 57.14% Mismatches: 5  
Query Match: 63.45% Indels: 0  
DB: 17 Gaps: 0

US-09-987-190-2 (1-30) x AZ929307 (1-519)

Qy 1 LyeTySerIeuProGluLeuAapTyGluPheserAlaThrGluProTyrllesergly 20  
Db 371 AAGGTACCTTGGCAAGATTAAGATTATTTCTCTACTGAACCATTAATCTCGGCT 430  
Qy 21 Glnlleasngluile\*\*\*TyrThr 28  
Db 431 CAAATCAACGAGTTGCACTACTCG 454

## RESULT 3

BQ087025

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BQ087025 464 bp mRNA linear EST 05-APR-2002  
Ct1\_9.A14.SP6 Ceratopteris Spore Library Ceratopteris richardii  
cDNA clone Ct1\_9.A14.5', mRNA sequence.  
BQ087025  
BQ087025.1 GI:20046226  
EST.  
Ceratopteris richardii.  
Ceratopteris richardii.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Filicophyta; Filicopsida; Filicales; Pteridaceae; Ceratopteris.  
Chatterjee, A., San Miguel, P., Stout, S.C., Banke, J. and Roux, S.J.  
Expressed sequence tags of cDNA clones from a C. richardii library  
Unpublished (2000)  
Contact: Roux SJ  
Section of Molecular Cell and Developmental Biology  
University of Texas  
Biology Building, Room 16, Austin, TX 78712, USA  
Tel: 512 471 4238  
Fax: 512 232 3402  
Email: sroux@uts.cc.utexas.edu  
Plate: Ct1\_9 row: A column: 14

Seq primer: SP6.  
Location/Qualifiers  
FEATURES  
source  
1..464  
/organism="Ceratomyxa richardii"  
/cultivar="Brogan"  
/db\_xref="taxon:49495"  
/clone="Cr1\_9\_A14"  
/tissue\_type="Ceratomyxa Spore Library"  
/cell\_type="Spore"  
/dev\_stage="20 hours after germination initiation"  
/note="Vector: pCMVSPORT5, EST sequence from cDNA library.  
cDNA library constructed from mRNA isolated from C.  
richardii spores that had developed for 20 hours after  
their germination had been initiated by white light."  
BASE COUNT 120 a 106 c 116 g 122 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 5.76e-05 Length: 464  
Score: 90.00 Matches: 16  
Percent Similarity: 81.48% Conservative: 6  
Best Local Similarity: 59.26% Mismatches: 5  
Query Match: 62.07% Indels: 0  
DB: 14 Gaps: 0  
US-09-987-190-2 (1-30) x BQ087025 (1-464)  
QY 2 TySerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGlyGln 21  
DB 154 TACTCTTACCCGAGCTAGCTTACGATTATGGGCGCTTAGAGCCGTACATCAGTGGCGAA 213  
QY 22 IleAsnGluIle\*\*TyrThr 28  
DB 214 ATTATGAGCTGCATCACAC 234  
RESULT 4  
LOCUS AI1319465 231 bp mRNA linear EST 18-DEC-1998  
DEFINITION b2a04nm.f1 Neurospora crassa morning cDNA library Neurospora crassa  
ACCESSION AI1319465  
CDNA clone b2a04nm 5', mRNA sequence.  
VERSION AI1319465.1 GI:4035447  
KEYWORDS EST.  
SOURCE Neurospora crassa.  
ORGANISM Neurospora crassa  
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
Sordariales; Sordariaceae; Neurospora.  
1 (bases 1 to 231)  
AUTHORS Zhu, H., Lai, H., Kupfer, D., Dunlap, J.C. and Roe, B.A.  
TITLE Two Neurospora crassa EST Databases  
JOURNAL Unpublished (1998)  
CONTACT: Bruce A. Roe, University of Oklahoma, broe@ou.edu  
Department of Chemistry and Biochemistry  
Advanced Center for Genome Technology, University of Oklahoma  
620 Parrington Oval, Norman, OK 73019, USA  
Tel: 405 325 4912  
Fax: 405 325 7762  
Email: broe@ou.edu  
We anticipate the future release of the cDNA clones to the Fungal  
Genetics Stock Center  
Seq primer: Universal Forward Primer  
High quality sequence stop: 50.  
location/Qualifiers  
FEATURES  
source  
1..231  
/organism="Neurospora crassa"  
/strain="bd, frq7 A"  
/db\_xref="taxon:5141"  
/clone="b2a04nm"  
/clone\_1lb="Neurospora crassa morning cDNA library"  
/tissue\_type="tissue harvested following 22hr growth in  
dark"  
/note="Vector: pBluescript SK-; Site\_1: XbaI; Site\_2:

ECORI; See: Bell-Pedersen, D., et al. PNAS 93:113096,1996.  
5' end of cDNA cloned into XbaI site of pBluescript; 3'  
end of cDNA cloned into EcoRI site of pBluescript"  
BASE COUNT 55 a 91 c 41 g 44 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 0.000304 Length: 231  
Score: 83.00 Matches: 15  
Percent Similarity: 77.78% Conservative: 6  
Best Local Similarity: 55.56% Mismatches: 6  
Query Match: 57.24% Indels: 0  
DB: 9 Gaps: 0  
US-09-987-190-2 (1-30) x AI1319465 (1-231)  
QY 2 TySerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGlyGln 21  
DB 65 TATTCCTTCCGAGCTACATATGCTTACGAGCCGTACATCAGCCAG 124  
QY 22 IleAsnGluIle\*\*TyrThr 28  
DB 125 ATCATGAGCTTACACAC 145  
RESULT 5  
LOCUS AI1320309 265 bp mRNA linear EST 18-DEC-1998  
DEFINITION c3b05nm.f1 Neurospora crassa morning cDNA library Neurospora crassa  
ACCESSION AI1320309  
CDNA clone c3b05nm 5', mRNA sequence.  
VERSION AI1320309.1 GI:4036291  
KEYWORDS EST.  
SOURCE Neurospora crassa.  
ORGANISM Neurospora crassa  
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
Sordariales; Sordariaceae; Neurospora.  
1 (bases 1 to 265)  
AUTHORS Zhu, H., Lai, H., Kupfer, D., Dunlap, J.C. and Roe, B.A.  
TITLE Two Neurospora crassa EST Databases  
JOURNAL Unpublished (1998)  
CONTACT: Bruce A. Roe, University of Oklahoma, broe@ou.edu  
Department of Chemistry and Biochemistry  
Advanced Center for Genome Technology, University of Oklahoma  
620 Parrington Oval, Norman, OK 73019, USA  
Tel: 405 325 4912  
Fax: 405 325 7762  
Email: broe@ou.edu  
We anticipate the future release of the cDNA clones to the Fungal  
Genetics Stock Center  
Seq primer: Universal Forward Primer  
High quality sequence stop: 187.  
location/Qualifiers  
FEATURES  
source  
1..265  
/organism="Neurospora crassa"  
/strain="bd, frq7 A"  
/db\_xref="taxon:5141"  
/clone="c3b05nm"  
/clone\_1lb="Neurospora crassa morning cDNA library"  
/tissue\_type="tissue harvested following 22hr growth in  
dark"  
/note="Vector: pBluescript SK-; Site\_1: XbaI; Site\_2:  
EcoRI; See: Bell-Pedersen, D., et al. PNAS 93:113096,1996.  
5' end of cDNA cloned into XbaI site of pBluescript; 3'  
end of cDNA cloned into EcoRI site of pBluescript"  
BASE COUNT 63 a 105 c 49 g 48 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 0.000372 Length: 265  
Score: 83.00 Matches: 15  
Percent Similarity: 77.78% Conservative: 6  
Best Local Similarity: 55.56% Mismatches: 6

Query Match:	57.24%	Indels:	0
DB:	9	Gaps:	0
US-09-987-190-2 (1-30) x A1320309 (1-265)			
OY	2	TyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGlyGln	21
DB	65	TATTCCTCTCCGACGTACCATATGCTCAATGCTCTGGACCTCACTACACCCAG	124
OY	22	IleAngIuIle**TyrThr	28
DB	125	ATCATGAGCTTCACACAGC	145
RESULT 6			
LOCUS	AW17655		
DEFINITION	h9g04nm.f1 Neurospora crassa morning cDNA library Neurospora crassa		
ACCESSION	AW17655		
VERSION	AW17655.1	GI:7606896	
KEYWORDS	EST.		
ORGANISM	Neurospora crassa.		
REFERENCE	Neurospora crassa		
AUTHORS	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;		
TITLE	Two Neurospora crassa EST databases		
JOURNAL	Unpublished (1998)		
COMMENT	Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu Department of Chemistry and Biochemistry Advanced Center for Genome Technology, University of Oklahoma 620 Parrington Oval, Norman, OK 73019, USA Tel: 405 325 4912 Fax: 405 325 7762 Email: broe@ou.edu		
FEATURES	We anticipate the future release of the cDNA clones to the Fungal Genetics Stock Center		
Source	Seq primer: Universal Forward Primer High quality sequence stop: 188. Location/Qualifiers 1..265 /organism="Neurospora crassa" /strain="bd, frq7 A" /db_xref="taxon:5141" /clone="h9g04nm" /clone_1ib="Neurospora crassa morning cDNA library" /tissue_type="tissue harvested following 22hr growth in dark" /note="Vector: pBluescript SK-; Site_1: XbaI; Site_2: EcoRI; See: Bell-Pedersen, D., et al. PNAS 93:13096,1996. 5' end of cDNA cloned into XbaI site of pBluescript; 3' end of cDNA cloned into EcoRI site of pBluescript"		
BASE COUNT	63 a	105 c	49 g
ORIGIN	48 t		
Alignment Scores:			
Pred. No.:	0.000372	Length:	265
Score:	83.00	Matches:	15
Percent Similarity:	77.78%	Conservative:	6
Best Local Similarity:	55.56%	Mismatches:	0
Query Match:	57.24%	Indels:	0
DB:	10	Gaps:	0
US-09-987-190-2 (1-30) x AW17655 (1-265)			
OY	2	TyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGlyGln	21
DB	65	TATTCCTCTCCGACGTACCATATGCTCAATGCTCTGGACCTCACTACACCCAG	124
OY	22	IleAngIuIle**TyrThr	28
DB	125	ATCATGAGCTTCACACAGC	145

LOCUS	267 bp	mRNA	linear	EST 16-DEC-1998
AI118805				
DEFINITION	a2d06nm.f1	Neurospora crassa morning cDNA library	Neurospora crassa	
ACCESSION	AI118805			
VERSION	AI118805.1	GI:4034746		
KEYWORDS	EST.			
SOURCE	Neurospora crassa.			
ORGANISM	Neurospora crassa			
REFERENCE	Eukaryote; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Sordariales; Sordariaceae; Neurospora.			
AUTHORS	1 (bases 1 to 267)			
TITLE	Zhu, H., Lai, H., Kupfer, D., Dunlap, J.C. and Roe, B.A.			
JOURNAL	Two Neurospora crassa EST Databases			
COMMENT	Unpublished (1998)			
	Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu			
	Department of Chemistry and Biochemistry			
	Advanced Center for Genome Technology, University of Oklahoma			
	620 Parrington Oval, Norman, OK 73019, USA			
	Tel: 405 325 4912			
	Fax: 405 325 7762			
	Email: broe@ou.edu			
	We anticipate the future release of the cDNA clones to the Fungal			
	Genetics Stock Center			
	Seq primer: Universal Forward Primer			
	High quality sequence stop: 192.			
FEATURES	Location/Qualifiers			
source	1..267			
	/organism="Neurospora crassa"			
	/strain="bd, frg7 A"			
	/db_xref="taxon:5141"			
	/clone="a2d06nm"			
	/clone_1lb="Neurospora crassa morning cDNA library"			
	/cbase_type="cbase harvested following 22hr growth in dark"			
	/note="Vector: pBluescript SK-; Site 1: XbaI; Site 2: EcoRI; See: Bell-Pedersen, D., et al. PNAS 93:13096,1996.			
	5' end of cDNA cloned into XbaI site of pBluescript; 3' end of cDNA cloned into EcoRI site of pBluescript"			
BASE COUNT	63 a 105 c 51 g 48 t			
ORIGIN				
Alignment Scores:				
Pred. No.:	0.000376	Length:	267	
Score:	83.00	Matches:	15	
Percent Similarity:	77.78%	Conservative:	6	
Best Local Similarity:	55.56%	Mismatch:	6	
Query Match:	57.24%	Indels:	0	
DB:	9	Gaps:	0	
US-09-987-190-2 (1-30) x AI118805 (1-267)				
Dy	2	TyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGlyGln	21	
Db	65	TATTCCTTCGCGACTACATATGCTCCTGAGCCCTACATCTCAGCCAG	124	
Gy	22	IleLeuGluIle***TyrThr	28	
Db	125	ATCATGAGCTTCACACAGC	145	
RESULT 8	AW723904	267 bp	mRNA	linear
LOCUS	AW723904			EST 19-APR-2000
DEFINITION	f2d01nm.f1	Neurospora crassa morning cDNA library	Neurospora crassa	
ACCESSION	AW723904			
VERSION	AW723904.1	GI:7618464		
KEYWORDS	EST.			
SOURCE	Neurospora crassa			
ORGANISM	Neurospora crassa			
	Eukaryote; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;			

REFERENCE 1 (bases 1 to 267)  
 AUTHORS Zhu, H., Lai, H., Kupfer, D., Dunlap, J.C., and Roe, B.A.  
 TITLE Two Neurospora crassa EST Databases  
 JOURNAL Unpublished (1998)  
 COMMENT Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu  
 Department of Chemistry and Biochemistry  
 Advanced Center for Genome Technology, University of Oklahoma  
 620 Parrington Oval, Norman, OK 73019, USA  
 Tel: 405 325 4912  
 Fax: 405 325 7762  
 Email: broe@ou.edu

FEATURES  
 source We anticipate the future release of the cDNA clones to the Fungal Genetics Stock Center  
 Seq primer: Universal Forward Primer  
 High quality sequence stop: 229.  
 Location/Qualifiers  
 1..267  
 /organism="Neurospora crassa"  
 /strain="bd, frq7 A"  
 /db\_xref="taxon:5141"  
 /clone="f2d01nm"  
 /clone\_1ib="Neurospora crassa morning cDNA library"  
 /tissue\_type="tissue harvested following 22hr growth in dark"  
 /note="Vector: pBluescript SK-; Site 1: XbaI; Site 2: EcoRI; See: Bell-Pedersen, D., et al. PNAS 93:13096, 1996.  
 5' end of cDNA cloned into XbaI site of pBluescript; 3' end of cDNA cloned into EcoRI site of pBluescript"

BASE COUNT 63 a 105 c 51 g 48 t  
 ORIGIN

## Alignment Scores:

Pred. No.:	0.000376	Length:	267
Score:	83.00	Matches:	15
Percent Similarity:	77.78%	Conservative:	6
Best Local Similarity:	55.56%	Mismatches:	6
Query Match:	57.24%	Indels:	0
DB:	10	Gaps:	0

US-09-987-190-2 (1-30) x AW723904 (1-267)

QY 2 TysSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGlyGln 21  
 |||||  
 Db 65 TATTCCTTCGAGACTACCATATGCTTCAATGCTCTGAGCCCTACATCTCAGCCAG 124  
 |||||

22 TleAsnGluIle\*\*TyrThr 28  
 |||||

125 ATCATGAGAGCTTCACCAAGC 145

RESULT 9 A1321638 269 bp mRNA linear EST 18-DEC-1998  
 LOCUS A1321638  
 DEFINITION elid5nm.f1 Neurospora crassa morning cDNA library Neurospora crassa  
 cDNA clone elid5nm 5', mRNA sequence.  
 A1321638  
 ACCESSION A1321638  
 VERSION A1321638.1 GI:4037620  
 KEYWORDS EST.

SOURCE Neurospora crassa.  
 ORGANISM Neurospora crassa  
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 Sordariales; Sordariaceae; Neurospora.

REFERENCE 1 (bases 1 to 269)  
 AUTHORS Zhu, H., Lai, H., Kupfer, D., Dunlap, J.C., and Roe, B.A.  
 TITLE Two Neurospora crassa EST Databases  
 JOURNAL Unpublished (1998)  
 COMMENT Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu  
 Department of Chemistry and Biochemistry  
 Advanced Center for Genome Technology, University of Oklahoma  
 620 Parrington Oval, Norman, OK 73019, USA  
 Tel: 405 325 4912  
 Fax: 405 325 7762  
 Email: broe@ou.edu

We anticipate the future release of the cDNA clones to the Fungal Genetics Stock Center  
 Seq primer: Universal Forward Primer  
 High quality sequence stop: 229.  
 Location/Qualifiers  
 1..269  
 /organism="Neurospora crassa"  
 /strain="bd, frq7 A"  
 /db\_xref="taxon:5141"  
 /clone="elid5nm"  
 /clone\_1ib="Neurospora crassa morning cDNA library"  
 /tissue\_type="tissue harvested following 22hr growth in dark"  
 /note="Vector: pBluescript SK-; Site 1: XbaI; Site 2: EcoRI; See: Bell-Pedersen, D., et al. PNAS 93:13096, 1996.  
 5' end of cDNA cloned into XbaI site of pBluescript; 3' end of cDNA cloned into EcoRI site of pBluescript"

BASE COUNT 63 a 106 c 52 g 48 t  
 ORIGIN

## Alignment Scores:

Pred. No.:	0.00038	Length:	269
Score:	83.00	Matches:	15
Percent Similarity:	77.78%	Conservative:	6
Best Local Similarity:	55.56%	Mismatches:	6
Query Match:	57.24%	Indels:	0
DB:	9	Gaps:	0

US-09-987-190-2 (1-30) x A1321638 (1-269)

QY 2 TysSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGlyGln 21  
 |||||  
 Db 65 TATTCCTTCGAGACTACCATATGCTTCAATGCTCTGAGCCCTACATCTCAGCCAG 124  
 |||||

QY 22 TleAsnGluIle\*\*TyrThr 28  
 |||||

Db 125 ATCATGAGAGCTTCACCAAGC 145

RESULT 10 A1320505  
 LOCUS A1320505

DEFINITION A1320505 336 bp mRNA linear EST 18-DEC-1998  
 csh03nm.f1 Neurospora crassa morning cDNA library Neurospora crassa  
 cDNA clone csh03nm 5', mRNA sequence.  
 A1320505  
 ACCESSION A1320505  
 VERSION A1320505.1 GI:4036487  
 KEYWORDS EST.

SOURCE Neurospora crassa.  
 ORGANISM Neurospora crassa  
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 Sordariales; Sordariaceae; Neurospora.

REFERENCE 1 (bases 1 to 336)  
 AUTHORS Zhu, H., Lai, H., Kupfer, D., Dunlap, J.C., and Roe, B.A.  
 TITLE Two Neurospora crassa EST Databases  
 JOURNAL Unpublished (1998)  
 COMMENT Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu  
 Department of Chemistry and Biochemistry  
 Advanced Center for Genome Technology, University of Oklahoma  
 620 Parrington Oval, Norman, OK 73019, USA  
 Tel: 405 325 4912  
 Fax: 405 325 7762  
 Email: broe@ou.edu

We anticipate the future release of the cDNA clones to the Fungal Genetics Stock Center  
 Seq primer: Universal Forward Primer  
 High quality sequence stop: 194.  
 Location/Qualifiers  
 1..336  
 /organism="Neurospora crassa"  
 /strain="bd, frq7 A"  
 /db\_xref="taxon:5141"  
 /clone="csh03nm"  
 /clone\_1ib="Neurospora crassa morning cDNA library"  
 /tissue\_type="tissue harvested following 22hr growth in dark"

dark"  
/note="Vector: pBluescript SK-; Site 1: XbaI; Site 2:  
EcoRI; See: Bell-Pedersen,D., et al. PNAS 93:13096,1996.  
5' end of cDNA cloned into XbaI site of pBluescript; 3'  
end of cDNA cloned into EcoRI site of pBluescript"

BASE COUNT 80 a 133 c 64 g 59 t

## ALIGNMENT Scores:

Pred. No.: 0.000527 Length: 336  
Score: 83.00 Matches: 15  
Percent Similarity: 77.78% Conservative: 6  
Best Local Similarity: 55.56% Mismatches: 6  
Query Match: 57.24% Indels: 0  
DB: 9 Gaps: 0

US-09-987-190-2 (1-30) x A1320505 (1-336)

QY 2 TySerLeuProGluLeuAspTYrGluPheSerAlaThrGluProTYrIleSerGlyIn 21

65 TATTCCTTCGCGAGCTACCATATGCTCTGAGAGCCCTACATCTACGCCAG 124

22 TleAsnGluIle\*\*TYrThr 28

Db 125 ATCATGAGAGCTTCACACAGC 145

## RESULT 11

A1318710 339 bp mRNA linear EST 18-DEC-1998

LOCUS a1b10m.f1 Neurospora crassa morning cDNA library Neurospora crassa

DEFINITION cDNA clone a1b10m 5', mRNA sequence.

ACCESSION A1318710

VERSION A1318710.1 GI:4034645

KEYWORDS EST.

SOURCE Neurospora crassa.

ORGANISM Neurospora crassa

REFERENCE Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;

AUTHORS Sordariales; Sordariaceae; Neurospora.

TITLE 1 (bases 1 to 339)

JOURNAL Two Neurospora crassa EST Databases

COMMENT Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu

Advanced Center for Genome Technology, University of Oklahoma

620 Parrington Oval, Norman, OK 73019, USA

Tel: 405 325 4912

Fax: 405 325 7762

Email: broe@ou.edu

We anticipate the future release of the cDNA clones to the Fungal

Genetics Stock Center

Seq primer: Universal Forward Primer

High quality sequence stop: 309.

Location/Qualifiers

1. .339

/organism="Neurospora crassa"

Best Local Similarity: 55.56% Mismatches: 6  
Query Match: 57.24% Indels: 0  
DB: 9 Gaps: 0

US-09-987-190-2 (1-30) x A1318710 (1-339)

QY 2 TySerLeuProGluLeuAspTYrGluPheSerAlaThrGluProTYrIleSerGlyIn 21

65 TATTCCTTCGCGAGCTACCATATGCTCTGAGAGCCCTACATCTACGCCAG 124

QY 22 TleAsnGluIle\*\*TYrThr 28

Db 125 ATCATGAGAGCTTCACACAGC 145

## RESULT 12

A1319577 339 bp mRNA linear EST 18-DEC-1998

LOCUS b3g06m.f1 Neurospora crassa morning cDNA library Neurospora crassa

DEFINITION cDNA clone b3g06m 5', mRNA sequence.

ACCESSION A1319577

VERSION A1319577.1 GI:4035559

KEYWORDS EST.

SOURCE Neurospora crassa.

ORGANISM Neurospora crassa

REFERENCE Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;

AUTHORS Sordariales; Sordariaceae; Neurospora.

TITLE 1 (bases 1 to 339)

JOURNAL Two Neurospora crassa EST Databases

COMMENT Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu

Advanced Center for Genome Technology, University of Oklahoma

620 Parrington Oval, Norman, OK 73019, USA

Tel: 405 325 4912

Fax: 405 325 7762

Email: broe@ou.edu

We anticipate the future release of the cDNA clones to the Fungal

Genetics Stock Center

Seq primer: Universal Forward Primer

High quality sequence stop: 314.

Location/Qualifiers

1. .339

/organism="Neurospora crassa"

/strain="bd, frq7 A"

/db\_xref="taxon:5141"

/clone="b3g06m"

/clone\_lib="Neurospora crassa morning cDNA library"

/tissue\_type="tissue harvested following 22hr growth in

dark"

/note="Vector: pBluescript SK-; Site 1: XbaI; Site 2:  
EcoRI; See: Bell-Pedersen,D., et al. PNAS 93:13096,1996.  
5' end of cDNA cloned into XbaI site of pBluescript; 3'  
end of cDNA cloned into EcoRI site of pBluescript"

BASE COUNT 82 a 133 c 65 g 59 t

ORIGIN

ALIGNMENT Scores:

Pred. No.: 0.000534 Length: 339

Score: 83.00 Matches: 15

Percent Similarity: 77.78% Conservative: 6

Best Local Similarity: 55.56% Mismatches: 6

Query Match: 57.24% Indels: 0

DB: 9 Gaps: 0

US-09-987-190-2 (1-30) x A1319577 (1-339)

QY 2 TySerLeuProGluLeuAspTYrGluPheSerAlaThrGluProTYrIleSerGlyIn 21

65 TATTCCTTCGCGAGCTACCATATGCTCTGAGAGCCCTACATCTACGCCAG 124

QY 22 TleAsnGluIle\*\*TYrThr 28

Db 125 ATCATGAGAGCTTCACACAGC 145





Email: broe@ou.edu  
We anticipate the future release of the cDNA clones to the Fungal  
Genetics Stock Center  
Seq primer: Universal Forward Primer  
High quality sequence stop: 394.  
Location/Qualifiers

## FEATURES

source

1. 401  
/organism="Neurospora crassa"  
/strain="bd, frq7 A"  
/db\_xref="taxon:5141"  
/clone="f8b07nm"  
/clone\_lib="Neurospora crassa morning cDNA library"  
/tissue\_type="tissue harvested following 22hr growth in  
dark"  
/note="Vector: pBluescript SK-; Site 1: XbaI; Site 2:  
EcoRI; See: Bell-Pedersen, D., et al. PNAS 93:13096,1996.  
5' end of cDNA cloned into XbaI site of pBluescript; 3'  
end of cDNA cloned into EcoRI site of pBluescript"

BASE COUNT

96 a 145 c 87 g 73 t

ORIGIN

## Alignment Scores:

Alignment Scores:  
Pred. No.: 0.000683 Length: 401  
Score: 83.00 Matches: 15  
Percent Similarity: 77.78% Conservative: 6  
Best Local Similarity: 55.56% Mismatches: 6  
Query Match: 57.24% Indels: 0  
DB: 10 Gaps: 0

US-09-987-190-2 (1-30) x AW724781 (1-401)

QY 2 TySerLeuProGluLeuAspTYrGluPheSerAlaThrGluProTYrIleSerGlyGln 21

DB 65 TATTCCCTTCGAGCTACCATATGCTTACCAATGCTCTGAGCCCTACATCTCAGCCAG 124

QY 22 IleAsnGluIle\*\*TYrThr 28

DB 125 ATCATGAGCTTCACACAGC 145

RESULT 16

LOCUS

AW715120 429 bp mRNA linear EST 19-APR-2000

DEFINITION g2c12nm.f1 Neurospora crassa morning cDNA library Neurospora crassa

ACCESSION AW715120 cDNA clone g2c12nm 5', mRNA sequence.

VERSION AW715120.1 GI:7604309

KEYWORDS

SOURCE

ORGANISM

Neurospora crassa.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

/tissue\_type="tissue harvested following 22hr growth in  
dark"  
/note="Vector: pBluescript SK-; Site 1: XbaI; Site 2:  
EcoRI; See: Bell-Pedersen, D., et al. PNAS 93:13096,1996.  
5' end of cDNA cloned into XbaI site of pBluescript; 3'  
end of cDNA cloned into EcoRI site of pBluescript"

BASE COUNT

98 a 154 c 95 g 82 t

ORIGIN

## Alignment Scores:

Alignment Scores:  
Pred. No.: 0.000754 Length: 429  
Score: 83.00 Matches: 15  
Percent Similarity: 77.78% Conservative: 6  
Best Local Similarity: 55.56% Mismatches: 6  
Query Match: 57.24% Indels: 0  
DB: 10 Gaps: 0

US-09-987-190-2 (1-30) x AW715120 (1-429)

QY 2 TySerLeuProGluLeuAspTYrGluPheSerAlaThrGluProTYrIleSerGlyGln 21

DB 65 TATTCCCTTCGAGCTACCATATGCTTACCAATGCTCTGAGCCCTACATCTCAGCCAG 124

QY 22 IleAsnGluIle\*\*TYrThr 28

DB 125 ATCATGAGCTTCACACAGC 145

RESULT 17

LOCUS

AW715952 429 bp mRNA linear EST 19-APR-2000

DEFINITION g8b12nm.f1 Neurospora crassa morning cDNA library Neurospora crassa

ACCESSION AW715952 cDNA clone g8b12nm 5', mRNA sequence.

VERSION AW715952.1 GI:7605157

KEYWORDS

SOURCE

ORGANISM

Neurospora crassa.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT



Alignment Scores:  
 Pred. No.: 0.00221  
 Score: 82.00  
 Percent Similarity: 75.00%  
 Best Local Similarity: 53.57%  
 Query Match: 56.55%  
 DB: 13  
 Gaps: 0

US-09-987-190-2 (1-30) x BM266232 (1-689)

QY 1 LyeTyrSerLeuProGluLeuAspTyrGluPheserAlaThrGluProTyrIleSerGly 20  
 DB 404 AAGTTACTCTTCGCGCATTCGCGTATGATGCGCTAGAACCTTACATCTCCGCGC 463

QY 21 GluIleAsnGluIle\*\*TyrThr 28  
 DB 464 CAATTATGCATTCACCATTC 487

RESULT 20  
 BM367690

LOCUS BM367690 466 bp mRNA linear EST 10-JAN-2002  
 DEFINITION NXLV\_053\_C03\_F\_NXLV (Nsf Xylem Late wood Vertical) Pinus taeda cDNA  
 clone NXLV\_053\_C03\_5', mRNA sequence.

ACCESSION BM367690  
 VERSION BM367690.1 GI:18111080  
 KEYWORDS EST.  
 SOURCE loblobly pine.  
 ORGANISM Pinus taeda

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.  
 1 (bases 1 to 466)

AUTHORS Sederoff, R.  
 TITLE Molecular Basis of Wood Formation in the Pine Megagenome  
 JOURNAL Unpublished (2000)

COMMENT Contact: Johnson, Arthur  
 North Carolina State University  
 Tel: 919 515 7800  
 Fax: 919 515 7801  
 Email: a.johnson@unc1ty.ncsu.edu  
 Seq primer: T3.

FEATURES  
 source Location/Qualifiers

1..466  
 /organism="Pinus taeda"  
 /strain="Coastal plain loblobly pine from North Carolina"  
 /db\_xref="taxon:3352"  
 /clone="NXLV\_053\_C03"  
 /clone\_1lb="NXLV (Nsf Xylem Late wood Vertical)"  
 /tissue\_type="primary xylem"  
 /dev\_stage="late wood"  
 /lab\_host="XLI-Blue"  
 /note="Vector: pTRIPLEX; Site 1: EcoRI; The library is from late (summer-August) wood, taken from below the crown of a 20 year old tree. The harvested xylem tissue was on the cusp between transitional and mature wood. NOTE: The sequences contain a 'cDNA adapter' between the EcoRI site and the start of the EST. The adapter sequence is 'AATCGGCATATGCGC'."

BASE COUNT 118 a 86 c 126 g 118 t 18 others

ORIGIN

Alignment Scores:  
 Pred. No.: 0.00269  
 Score: 80.00  
 Percent Similarity: 76.92%  
 Best Local Similarity: 57.69%  
 Query Match: 55.17%  
 DB: 13  
 Gaps: 0

US-09-987-190-2 (1-30) x BM367690 (1-466)

QY 2 TyrSerLeuProGluLeuAspTyrGluPheserAlaThrGluProTyrIleSerGly 21  
 DB 204 TTATGTTACCGAGCTTCATAGATACAGCGCATTCGAAACCGCTTATTATGATGAG 263

QY 22 IleAsnGluIle\*\*Tyr 27  
 DB 264 ATTATGAGCTGCACAC 281

RESULT 21  
 AW733168

LOCUS AW733168 306 bp mRNA linear EST 03-DEC-2001  
 DEFINITION sk70h06.y1 Gm-c1016 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-c1016-9492 5' similar to SW:SDM\_PEA P27084 SUPEROXIDE DISMUTASE [MN] PRECURSOR ;, mRNA sequence.

ACCESSION AW733168  
 VERSION AW733168.1 GI:7638902  
 KEYWORDS EST.  
 SOURCE soybean.  
 ORGANISM Glycine max

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosida I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.  
 1 (bases 1 to 306)

AUTHORS Shoemaker, R., Kain, P., Vodkin, L., Erpelting, J., Corryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kueba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritzer, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterson, R. and Wilson, R.  
 TITLE Public Soybean EST Project  
 JOURNAL Unpublished (1999)

COMMENT Contact: Shoemaker R/Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810

Email: est@watson.wustl.edu  
 This clone is available through: Resgen, Invitrogen Corp. 2130 South Memorial Parkway Huntville, AL 35801 For further information call: (800)-533-4363 or contact via email: cduoresgen.com  
 Insert Length: 1240 Std Error: 0.00  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 305.

FEATURES  
 source Location/Qualifiers

1..306  
 /organism="Glycine max"  
 /db\_xref="taxon:3847"  
 /clone="GENOME SYSTEMS CLONE ID: Gm-c1016-9492"  
 /clone\_1lb="Gm-c1016"  
 /tissue\_type="immature flowers of field grown plants"  
 /dev\_stage="XLI0-Gold"  
 /lab\_host="XLI0-Gold"  
 /note="Vector: pBluescript II XR; Site 1: EcoRI; Site 2: XhoI. This cDNA library was constructed from mRNA isolated from immature flowers of field grown plants. The cDNA library was prepared using the Stratagene pBluescript II XR library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly (dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into XLI0-Gold host cells. This library was constructed by Dr. Randy Shoemaker and Dr. John Erpelting."

BASE COUNT 72 a 102 c 72 g 60 t

ORIGIN

Alignment Scores:  
 Pred. No.: 0.00213  
 Score: 79.00  
 Percent Similarity: 80.77%  
 Best Local Similarity: 50.00%  
 Query Match: 54.48%  
 DB: 13  
 Gaps: 0





REFERENCE	TITLE	COMMENT	FEATURES
1 (bases 1 to 448)	Public Soybean EST Project	Unpublished (1999)	Location/Qualifiers
Shoemaker, R., Keim, P., Vodkin, L., Erpelidg, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kuabab, T., Martin, T., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.	Public Soybean EST Project	Unpublished (1999)	1. 448
Public Soybean EST Project	Public Soybean EST Project	Public Soybean EST Project	/db xref="taxon:3847"
Washington University School of Medicine	Washington University School of Medicine	Washington University School of Medicine	/clone="GENOME SYSTEMS CLONE ID: Gm-cl004-4579"
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA	/clone_1id="Gm-cl004"
Tel: 314 286 1800	Tel: 314 286 1800	Tel: 314 286 1800	/clone_2id="Gm-cl004"
Fax: 314 286 1810	Fax: 314 286 1810	Fax: 314 286 1810	/clone_3id="Gm-cl004"
Email: est@watson.wustl.edu	Email: est@watson.wustl.edu	Email: est@watson.wustl.edu	/clone_4id="Gm-cl004"
Possible reversed clone: similarity on wrong strand This clone is available through: Resgen, Invitrogen Corp, 2130 South Memorial Parkway Huntville, AL 35801 For further information call: (800) 533-4363 or contact via email: ccl@resgen.com	Possible reversed clone: similarity on wrong strand This clone is available through: Resgen, Invitrogen Corp, 2130 South Memorial Parkway Huntville, AL 35801 For further information call: (800) 533-4363 or contact via email: ccl@resgen.com	Possible reversed clone: similarity on wrong strand This clone is available through: Resgen, Invitrogen Corp, 2130 South Memorial Parkway Huntville, AL 35801 For further information call: (800) 533-4363 or contact via email: ccl@resgen.com	/clone_5id="Gm-cl004"
Insert Length: 1069 Std Error: 0.00	Insert Length: 1069 Std Error: 0.00	Insert Length: 1069 Std Error: 0.00	/clone_6id="Gm-cl004"
Seq primer: -40RP from Gibco	Seq primer: -40RP from Gibco	Seq primer: -40RP from Gibco	/clone_7id="Gm-cl004"
High quality sequence stop: 124	High quality sequence stop: 124	High quality sequence stop: 124	/clone_8id="Gm-cl004"
POLYA=No	POLYA=No	POLYA=No	/clone_9id="Gm-cl004"
Location/Qualifiers	Location/Qualifiers	Location/Qualifiers	/clone_10id="Gm-cl004"
1. 448	1. 448	1. 448	/clone_11id="Gm-cl004"
/db xref="taxon:3847"	/db xref="taxon:3847"	/db xref="taxon:3847"	/clone_12id="Gm-cl004"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl004-4579"	/clone="GENOME SYSTEMS CLONE ID: Gm-cl004-4579"	/clone="GENOME SYSTEMS CLONE ID: Gm-cl004-4579"	/clone_13id="Gm-cl004"
/clone_1id="Gm-cl004"	/clone_1id="Gm-cl004"	/clone_1id="Gm-cl004"	/clone_14id="Gm-cl004"
/clone_2id="Gm-cl004"	/clone_2id="Gm-cl004"	/clone_2id="Gm-cl004"	/clone_15id="Gm-cl004"
/clone_3id="Gm-cl004"	/clone_3id="Gm-cl004"	/clone_3id="Gm-cl004"	/clone_16id="Gm-cl004"
/clone_4id="Gm-cl004"	/clone_4id="Gm-cl004"	/clone_4id="Gm-cl004"	/clone_17id="Gm-cl004"
/clone_5id="Gm-cl004"	/clone_5id="Gm-cl004"	/clone_5id="Gm-cl004"	/clone_18id="Gm-cl004"
/clone_6id="Gm-cl004"	/clone_6id="Gm-cl004"	/clone_6id="Gm-cl004"	/clone_19id="Gm-cl004"
/clone_7id="Gm-cl004"	/clone_7id="Gm-cl004"	/clone_7id="Gm-cl004"	/clone_20id="Gm-cl004"
/clone_8id="Gm-cl004"	/clone_8id="Gm-cl004"	/clone_8id="Gm-cl004"	/clone_21id="Gm-cl004"
/clone_9id="Gm-cl004"	/clone_9id="Gm-cl004"	/clone_9id="Gm-cl004"	/clone_22id="Gm-cl004"
/clone_10id="Gm-cl004"	/clone_10id="Gm-cl004"	/clone_10id="Gm-cl004"	/clone_23id="Gm-cl004"
/clone_11id="Gm-cl004"	/clone_11id="Gm-cl004"	/clone_11id="Gm-cl004"	/clone_24id="Gm-cl004"
/clone_12id="Gm-cl004"	/clone_12id="Gm-cl004"	/clone_12id="Gm-cl004"	/clone_25id="Gm-cl004"
/clone_13id="Gm-cl004"	/clone_13id="Gm-cl004"	/clone_13id="Gm-cl004"	/clone_26id="Gm-cl004"
/clone_14id="Gm-cl004"	/clone_14id="Gm-cl004"	/clone_14id="Gm-cl004"	/clone_27id="Gm-cl004"
/clone_15id="Gm-cl004"	/clone_15id="Gm-cl004"	/clone_15id="Gm-cl004"	/clone_28id="Gm-cl004"
/clone_16id="Gm-cl004"	/clone_16id="Gm-cl004"	/clone_16id="Gm-cl004"	/clone_29id="Gm-cl004"
/clone_17id="Gm-cl004"	/clone_17id="Gm-cl004"	/clone_17id="Gm-cl004"	/clone_30id="Gm-cl004"
/clone_18id="Gm-cl004"	/clone_18id="Gm-cl004"	/clone_18id="Gm-cl004"	/clone_31id="Gm-cl004"
/clone_19id="Gm-cl004"	/clone_19id="Gm-cl004"	/clone_19id="Gm-cl004"	/clone_32id="Gm-cl004"
/clone_20id="Gm-cl004"	/clone_20id="Gm-cl004"	/clone_20id="Gm-cl004"	/clone_33id="Gm-cl004"
/clone_21id="Gm-cl004"	/clone_21id="Gm-cl004"	/clone_21id="Gm-cl004"	/clone_34id="Gm-cl004"
/clone_22id="Gm-cl004"	/clone_22id="Gm-cl004"	/clone_22id="Gm-cl004"	/clone_35id="Gm-cl004"
/clone_23id="Gm-cl004"	/clone_23id="Gm-cl004"	/clone_23id="Gm-cl004"	/clone_36id="Gm-cl004"
/clone_24id="Gm-cl004"	/clone_24id="Gm-cl004"	/clone_24id="Gm-cl004"	/clone_37id="Gm-cl004"
/clone_25id="Gm-cl004"	/clone_25id="Gm-cl004"	/clone_25id="Gm-cl004"	/clone_38id="Gm-cl004"
/clone_26id="Gm-cl004"	/clone_26id="Gm-cl004"	/clone_26id="Gm-cl004"	/clone_39id="Gm-cl004"
/clone_27id="Gm-cl004"	/clone_27id="Gm-cl004"	/clone_27id="Gm-cl004"	/clone_40id="Gm-cl004"
/clone_28id="Gm-cl004"	/clone_28id="Gm-cl004"	/clone_28id="Gm-cl004"	/clone_41id="Gm-cl004"
/clone_29id="Gm-cl004"	/clone_29id="Gm-cl004"	/clone_29id="Gm-cl004"	/clone_42id="Gm-cl004"
/clone_30id="Gm-cl004"	/clone_30id="Gm-cl004"	/clone_30id="Gm-cl004"	/clone_43id="Gm-cl004"
/clone_31id="Gm-cl004"	/clone_31id="Gm-cl004"	/clone_31id="Gm-cl004"	/clone_44id="Gm-cl004"
/clone_32id="Gm-cl004"	/clone_32id="Gm		

Score:	79.00	Matches:	13
Percent Similarity:	80.77%	Conservative:	8
Best Local Similarity:	50.00%	Mismatches:	5
Query Matchn:	54.48%	Indels:	0
DB:	9	Gaps:	0

US-09-987-190-2 (1-30) × A1460811 (1-448)

Qy	2	TyrsrLeuPProGluLeuAspTyrGluPhseSerAlaThrGluProTyrIleSerGlycIn 21
Db	173	TACAGCTCCCGATCTGATTAACGACATATGCGCTCTGAGCCAGCCATCATCGAGCGCGAA 232
Qy	22	Ileangluile**Tyr 27
Db	233	ATCATGACGTGCACAC 250

RESULT 26

A1460823

LOCUS

DEFINITION

448 bp mRNA linear EST 28-NOV-2001

ba69f10.y1 Gm-cl1004 Glycine max cDNA clone GENOME SYSTEMS CLON ID: Gm-cl1004-4580 5' similar to SW:SODM\_PEA P27084 SUPEROXIDE DISMUTASE.

PRECUSOR ; mRNA sequence.

A1460823

VERSION

A1460823.1 GI:4313704

KEYWORDS

EST.

SOURCE

soybean.

ORGANISM

Glycine max

REFERENCE

1 (bases 1 to 448)

Shoemaker,R., Keim,P., Vodkin,L., Erpellding,J., Coryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Strepeou,M., Theising,B., Allen,N., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.

Public Soybean EST Project

Unpublished (1999)

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: east@watson.wustl.edu

Possible reversed clone: similarity on wrong strand. This clone is available through: Regen, Invitrogen Corp, 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact via email: ccu@regen.com

Insert Length: 1069 Std Error: 0.00

Seq primer: -40RP from Glibco

High quality sequence stop: 284

POLYA=No.

location/Qualifiers

1. .448

/organism="Glycine max"

/db\_xref="taxon:1847"

/clone="GENOME SYSTEMS CLONE ID: Gm-cl1004-4580"

/c1one\_id="Gm-cl1004"

/lab\_type="root"

/tax host="X110-Gold"

/note="vector: pluescript II XR; site 1: EcoRI; site 2: XhoI; Root cDNA. The mRNA was isolated from entire roots of 8 day old 'Williams' seedlings which were propagated on paper towels with distilled water. Stragene's cDNA Synthesis Kit (catalog #200401) was used to synthesize the cDNA. First- strand synthesis was performed with 5-methyl dCpP, hence the ligated cDNA is hemimethylated. Stragene's first-strand synthesis primer was used [GACAGGAGAGAGAGAGAACTAGTCTGCAG(T)-18]. After second-strand synthesis, the cDNA ends were 'polished'

FEATURES

source

TITLE

JOURNAL

COMMENT









REFERENCE 1 (bases 1 to 552)  
 soybean.  
 Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Corvelli, V., Khanna  
 A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,  
 Wylie, T., Underwood, K., Stepien, M., Theising, B., Allen, M., Bowers  
 Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk  
 R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann  
 R., Waterston, R. and Wilson, R.  
 Public Soybean EST Project  
 Unpublished (1999)  
 CONTACT: Shoemaker R./Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available through: ResGen, Invitrogen Corp. 2130  
 South Memorial Parkway Huntville, AL 35801 For further information  
 call: (800)-533-4363 or contact via email: ccu@resgen.com  
 Insert Length: 1152 Std Error: 0.00  
 High quality sequence stop: 414.  
 Location/Qualifiers  
 source 1..552  
 /organism="Glycine max"  
 /db\_xref="taxon:3847"  
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl039-480"  
 /clone\_1ib="Gm-cl039"  
 /tissue\_type="whole seedling without cotyledons"  
 /lab\_host="DH10B"  
 /note="Vector: pBluescript II SK+, Site 1: EcoRI; Site 2:  
 XhoI; This cDNA library was constructed from mRNA isolated  
 from 2 week old seedlings with the cotyledons removed at  
 the time of harvest. The seedlings for the cultivar Ogden  
 were grown in a growth chamber using germination paper.  
 Complementary DNA was synthesized from mRNA using a primer  
 consisting of a poly(dT) sequence with a XhoI restriction  
 site. EcoRI adapters were ligated to the blunt-ended cDNA  
 fragments followed by XhoI digestion. The cDNA fragments  
 were directionally cloned into the EcoRI-XhoI restriction  
 site of the pBluescript vector. The ligated cDNA fragments  
 were transformed into DH10B host cells (Gibco BRL). This  
 library was constructed by Dr. Randy Shoemaker."  
 BASE COUNT 130 a 159 c 145 g 117 t 1 others  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 0.00506 Length: 552  
 Score: 79.00 Matches: 13  
 Percent Similarity: 80.77% Conservative: 8  
 Best Local Similarity: 50.00% Mismatches: 5  
 Query Match: 54.48% Indels: 0  
 DB: 10 Gaps: 0  
 US-09-987-190-2 (1-30) x BR211438 (1-552)  
 QY 2 TysSerLeuProGluLeuAspTYRGluPhseSerAlaThrGluProTYRilesGlyGln 21  
 Db 169 TACACGCTCCCGATCTGCACTATGCGCTCTGAGCCAGCCATCAGCGGTGAA 228  
 QY 22 IleAsnGluIle\*\*TYR 27  
 Db 229 ATCATGCACTGCACAC 246  
 RESULT 31  
 LOCUS B1973216 564 bp mRNA linear EST 30-NOV-2001  
 DEFINITION sai55d1.y1 Gm-cl065 Glycine max cDNA clone GENOME SYSTEMS CLONE  
 ID: Gm-cl065-7845 5' similar to SW:SDOM\_PBA P27084 SUPEROXIDE  
 DISMUTASE [MN], MITOCHONDRIAL PRECURSOR ;, mRNA sequence.  
 ACCESSION B1973216  
 VERSION B1973216.1 GI:16347621  
 KEYWORDS EST.

SOURCE  
 ORGANISM  
 Glycine max  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.  
 1 (bases 1 to 564)  
 Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Corvelli, V., Khanna  
 A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,  
 Wylie, T., Underwood, K., Stepien, M., Theising, B., Allen, M., Bowers  
 Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk  
 R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann  
 R., Waterston, R. and Wilson, R.  
 Public Soybean EST Project  
 Unpublished (1999)  
 CONTACT: Shoemaker R./Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available through: ResGen, Invitrogen Corp. 2130  
 South Memorial Parkway Huntville, AL 35801 For further information  
 call: (800)-533-4363 or contact via email: ccu@resgen.com  
 High quality sequence stop: 421.  
 Location/Qualifiers  
 source 1..564  
 /organism="Glycine max"  
 /db\_xref="taxon:3847"  
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl065-7845"  
 /clone\_1ib="Gm-cl065"  
 /tissue\_type="germinating shoots"  
 /lab\_host="DH10B"  
 /note="Vector: pBluescript II SK+, Site 1: EcoRI; Site 2:  
 XhoI; The cDNA library was constructed from mRNA isolated  
 from germinating shoots of the cultivar Williams. The seeds  
 were allowed to germinate for 24 hours prior to being  
 cold stressed for 2 days at 4C. Complementary DNA was  
 synthesized from mRNA using a primer consisting of a  
 poly(dT) sequence with a XhoI restriction site. EcoRI  
 adapters were ligated to the blunt-ended cDNA fragments  
 followed by XhoI digestion. The cDNA fragments were  
 directionally cloned into the EcoRI-XhoI restriction site  
 of the pBluescript vector. The ligated cDNA fragments were  
 transformed into DH10B host cells (GibcoBRL). This library  
 was constructed in the laboratory of Dr. Randy  
 Shoemaker."  
 BASE COUNT 138 a 157 c 146 g 123 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 0.00522 Length: 564  
 Score: 79.00 Matches: 13  
 Percent Similarity: 80.77% Conservative: 8  
 Best Local Similarity: 50.00% Mismatches: 5  
 Query Match: 54.48% Indels: 0  
 DB: 13 Gaps: 0  
 US-09-987-190-2 (1-30) x B1973216 (1-564)  
 QY 2 TysSerLeuProGluLeuAspTYRGluPhseSerAlaThrGluProTYRilesGlyGln 21  
 Db 157 TACACGCTCCCGATCTGCACTATGCGCTCTGAGCCAGCCATCAGCGGTGAA 216  
 QY 22 IleAsnGluIle\*\*TYR 27  
 Db 217 ATCATGCACTGCACAC 234  
 RESULT 32  
 LOCUS BMS21591 571 bp mRNA linear EST 15-FEB-2002  
 DEFINITION saks59g10.y1 Gm-cl036 Glycine max cDNA clone SOYBEAN CLONE ID:

Gm-c1036-7603 5' similar to SW:SODM\_PEA P27084 SUPEROXIDE DISMUTASE [MN], MITOCHONDRIAL PRECURSOR ;, mRNA sequence.

ACCESSION BM521591  
VERSION BM521591.1 GI:18692743  
KEYWORDS EST.  
SOURCE soybean.  
ORGANISM Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

REFERENCE 1 (bases 1 to 571)

Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shih,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.  
Public Soybean EST Project  
Unpublished (1999)  
Contact: Shoemaker R/Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available through: Resgen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact: ccut@resgen.com web site: www.resgen.com

Seq primer: -40RP from Gibco  
High quality sequence stop: 428.

FEATURES

source

1..571  
/organism="Glycine max"  
/db\_xref="taxon:3847"  
/clone="SOYBEAN CLONE ID: Gm-c1036-7603"  
/clone\_1ib="Gm-c1036"  
/tissue\_type="somatic embryo cultured on MSD 20"  
/lab\_host="DH10B"  
/note="Vector: pSPORT1; Site 1: NotI; Site 2: SalI; This cDNA library was constructed from mRNA isolated from somatic embryos (age ranging from 2 months to 9 months) cultured on MSD 20. The library was prepared using the Life Technologies superscript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly (dT) sequence with a NotI restriction site. SalI linker adapters were ligated to the blunt-ended cDNA fragments followed by NotI digestion. The cDNA fragments were directionally cloned into the NotI-SalI restriction site of the pSPORT1 vector. The ligated cDNA fragments were transformed into B.coli Electromax DH10B host cells. This library was constructed in the laboratory of Dr. Lila Vodkin by Ann Khanna e-mail: l-vodkin@uiuc.edu"  
Urbana-Champaign

BASE COUNT 139 a 157 c 147 g 128 t

ORIGIN

Alignment Scores:  
Pred. No.: 0.00532 Length: 571  
Score: 79.00 Matches: 13  
Percent Similarity: 80.77% Conservative: 8  
Best Local Similarity: 50.00% Mismatches: 5  
Query Match: 54.48% Indels: 0  
DB: 13 Gaps: 0

US-09-987-190-2 (1-30) x BM521591 (1-571)

QY 2 TySeLenProGluLeuAapTyrgLupheSezAlaThrGluProTryIleSeGlyGln 21

DB 155 TACACGCTCCCGCATCTGGATTACGACTATGCGCTCTGAGCAGCAGCATCAGCGGCA 214

QY 22 llaenGluIle\*\*\*Tyr 27  
DB 215 ATCATGCACTGCACCA 232

RESULT 33  
LOCUS B0094786

DEFINITION B0094786 578 bp mRNA linear EST 08-APR-2002  
san51905.y1 Gm-c1052 Glycine max cDNA clone SOYBEAN CLONE ID: Gm-c1052-3370 5' similar to SW:SODM\_PEA P27084 SUPEROXIDE DISMUTASE [MN], MITOCHONDRIAL PRECURSOR ;, mRNA sequence.

ACCESSION B0094786  
VERSION B0094786.1 GI:20076003  
KEYWORDS EST.  
SOURCE soybean.  
ORGANISM Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

REFERENCE 1 (bases 1 to 578)

Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shih,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.  
Public Soybean EST Project  
Unpublished (1999)  
Contact: Shoemaker R/Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available through: Resgen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact: ccut@resgen.com web site: www.resgen.com

Seq primer: -40RP from Gibco  
High quality sequence stop: 427.

FEATURES

source

1..578  
/organism="Glycine max"  
/db\_xref="taxon:3847"  
/clone="SOYBEAN CLONE ID: Gm-c1052-3370"  
/clone\_1ib="Gm-c1052"  
/tissue\_type="whole seedlings of greenhouse grown plants"  
/dev\_stage="1 week old"  
/lab\_host="DH10B"  
/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI. The Harosoy NIL was constructed and seed was provided by Dr. J. Specht, University of Nebraska (Shoemaker and Specht, 1995). The cDNA library was constructed from mRNA isolated from whole seedlings of 1 week old greenhouse grown plants. Complementary DNA was synthesized from mRNA using a primer consisting of a 3' poly(dT) sequence with a XhoI restriction site and a 3' anchor. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). The library was constructed in cooperation with Dr. Paul Keim's laboratory at Northern Arizona University."

BASE COUNT 144 a 160 c 148 g 126 t

ORIGIN

Alignment Scores:  
Pred. No.: 0.00541 Length: 578  
Score: 79.00 Matches: 13  
Percent Similarity: 80.77% Conservative: 8



```

/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl053-1609"
/clone_lib="Gm-cl053"
/tissue_type="Whole seedling, 3 week old, greenhouse
grown"
/lab_host="DH10B"
/notes="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
XhoI; The Harosoy N1L was constructed and seed was
provided by Dr. J. Specht, University of Nebraska

```

Public Soybean EST Project  
Washington University School of Medicine  
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available through: ResGen, Invitrogen Corp. 2130  
South Memorial Parkway Hunttsville, AL 35801 For further information  
call: (800)-533-4363 or contact via email: ccu@resgen.com  
Insert Length: 1056 Std Error: 0.00  
High quality sequence stop: 413.  
Location/Qualifiers  
1..598  
/organism="Glycine max"  
FEATURES  
source

```

/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1028-997"
/clone_1b="Gm-c1028"
/tissue_type="roots of 'Superpod' plants"
/lab_host="DH10B"
/notes="Vector: pBluescript II Xr, Site_1: EcoRI, Site_2:
XhoI; The mRNA was isolated from roots of Glycine max
'Superpod' plants generously donated by Dr. Gary Stacey.
The seedlings were inoculated with Bradyrhizobium
japonicum, strain USDA110 prior to harvest. Stratagene's
cDNA synthesis kit (catalog number 200401) was used to
synthesize the cDNA. First-strand synthesis was performed
with 5-methyl dCTP, hence the ligated cDNA was
hemimethylated. A modification of Stratagene's
first-strand synthesis primer was used. An 'anchor'
nucleotide (V=A, C, or G) was added to the 3' end of the
primer (GAGAGAGAGAGAGAGAGAGAACTACTCTCGAG(T18V)) to anchor
the primer at the 5' end of the poly(A) tract. After
second-strand synthesis, the cDNA ends were filled in with
cloned Pfu DNA polymerase, ligated to EcoRI adapters and
subsequently phosphorylated. The XhoI site within the
first-strand synthesis primer was then restricted by
digestion with XhoI; all XhoI sites in the cDNA would be
protected by their hemimethylated status. The cDNA
constructs were size-fractionated with a 500bp cutoff,
using GibcoBRL Life Technologies' cDNA Size Fractionation
column. The column eluent was then ligated into
Stratagene's pBluescript II Xr Predigested vector
(pBluescript II SK(+)) that has been digested with EcoRI
and XhoI, and phosphorylated by Stratagene). Both the
white and blue colonies appear to contain recombinant
plasmids with cDNA inserts, based on size (n=25). This
library was constructed by Dr. Paul Keim and Dr. Virginia
Coryell."

BASE COUNT      147 a      166 c      151 g      132 t      2 others
ORIGIN

Alignment Scores:
Pred. No.:      0.00569      Length:      598
Score:          79.00      Matches:      13
Percent Similarity: 80.77%      Conservative: 8
Best Local Similarity: 50.00%      Mismatches: 5
Query Match:    54.48%      Indels:      0
DB:             10      Gaps:      0

US-09-987-190-2 (1-30) x AM234447 (1-598)

QY      2  TysSerLeuProGluLeuAspTYrGluPheSerAlaThrGluProTYrIleSerGlyGln 21
      173 TACACGCTCCCCGATCTGCAATTACGACTATGGCGCTCTGGAGCAGCACCATCAGCGGTGAA 232
QY      22 TleAsnGluIle***TYr 27
      233 ATCATGACGCTGCACAC 250

RESULT 38
LOCUS      BE330334      670 bp      mRNA      linear      EST 04-DEC-2001
DEFINITION B077004.Y1 Gm-c1040 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
Gm-c1040-1664 5', similar to SW:SDOM_PEA P27084 SUPEROXIDE DISMUTASE
[MN] PRECURSOR ;, mRNA sequence.
ACCESSION BE330334
VERSION BE330334.1 GI:9204110
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosoids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 670)
REFERENCE 1 (bases 1 to 670)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna
,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,

```

```

TITLE
JOURNAL
COMMENT
Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
,Y., Person,B., Waller,T., Gibbons,M., Page,D., Harvey,N., Schurk
,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
This clone is available through: Reggen, Invitrogen Corp. 2130
South Memorial Parkway Huntville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccu@reggen.com
Insert Length: 1066 Std Error: 0.00
High quality sequence stop: 470.
Location/Qualifiers
1..670
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1040-1664"
/clone_1b="Gm-c1040"
/tissue_type="Hypocotyl and Plumule, germinating seeds"
/lab_host="DH10B"
/notes="Vector: pT7T3Pac (Pharmacia); Site_1: EcoRI;
Site_2: NotI. This cDNA library was constructed from mRNA
isolated from hypocotyl and plumule tissues of seeds
germinated for three days of the cultivar Williams 82.
Complementary DNA was synthesized from mRNA using a primer
consisting of a poly(dT) sequence with a NotI restriction
site. EcoRI adapters were ligated to the blunt-ended cDNA
fragments followed by digestions with EcoRI and NotI. The
cDNA fragments were directionally cloned into the
EcoRI-NotI restriction site of the pT7T3-Pac vector. The
ligated cDNA fragments were transformed into DH10B host
cells (Gibco BRL). This library was constructed by Dr.
Randy Shoemaker."

BASE COUNT      162 a      177 c      178 g      151 t      2 others
ORIGIN

Alignment Scores:
Pred. No.:      0.00672      Length:      670
Score:          79.00      Matches:      13
Percent Similarity: 80.77%      Conservative: 8
Best Local Similarity: 50.00%      Mismatches: 5
Query Match:    54.48%      Indels:      0
DB:             10      Gaps:      0

US-09-987-190-2 (1-30) x BE330334 (1-670)

QY      2  TysSerLeuProGluLeuAspTYrGluPheSerAlaThrGluProTYrIleSerGlyGln 21
      173 TACACGCTCCCCGATCTGCAATTACGACTATGGCGCTCTGGAGCAGCACCATCAGCGGTGAA 232
QY      22 TleAsnGluIle***TYr 27
      233 ATCATGACGCTGCACAC 250

RESULT 39
LOCUS      BI263359      675 bp      mRNA      linear      EST 18-JUL-2001
DEFINITION NF089FP09P1F1079 Phosphate starved leaf Medicago truncatula cDNA
clone NF089FP09PL 5', mRNA sequence.
ACCESSION BI263359
VERSION BI263359.1 GI:14864520
KEYWORDS EST.
SOURCE barrel medic.
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosoids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;

```

REFERENCE 1 (bases 1 to 675)  
 AUTHORS Liu, J., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J., Flores, H.R., Innan, J.T., Weller, J.W., May, G.D., and Harrison, M.J.  
 TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation Medicago truncatula phosphate-starved leaf library  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Harrison MJ  
 Plant Biology Division  
 The Samuel Roberts Noble Foundation  
 2510 Sam Noble Parkway, Ardmore, OK 73402, USA  
 Tel: 580 221 7325  
 Fax: 580 221 7380  
 Email: mjharrison@noble.org  
 Insert Length: 675 Std Error: 0.00  
 Plate: 089 row: F column: 09  
 Seq primer: TCACACGAGGAACAGCTAGAC.  
 FEATURES  
 source  
 1.675  
 /organism="Medicago truncatula"  
 /db\_xref="taxon:3880"  
 /clone="NF089F09PL"  
 /tissue\_type="phosphate starved leaf"  
 /dev\_stage="trifoliolate"  
 /note="Vector: Lambda Zap; At the trifoliolate stage, M. truncatula plants were transplanted to phosphate-free sand and grown for a further 30 days. During this 30 day period, the plants were fertilized twice weekly with 1/2 Hoaglands solution containing only 200m potassium phosphate. RNA was prepared from above ground tissues."  
 BASE COUNT 134 a 240 c 146 g 153 t 2 others  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 0.00679 Length: 675  
 Score: 79.00 Matches: 14  
 Percent Similarity: 78.57% Conservative: 8  
 Best Local Similarity: 50.00% Mismatches: 6  
 Query Match: 54.48% Indels: 0  
 DB: 13 Gaps: 0  
 US-09-987-190-2 (1-30) x BI263359 (1-675)  
 QY 1 lvsTyrSerleuProGluLeuAspTYRGIuPheSerAlaThrGluProTYRleSergIy 20  
 172 AAGGCGACCTCCCGATCTTCTTACGACTATGCGCCTCTTGAGCTCACAATCCGAC 231  
 21 GlnIleAsnGluIle\*\*TYRthr 28  
 DB 232 CAGATCATGAGCTTCACACATCC 255  
 RESULT 40  
 BI942249  
 LOCUS BI942249 683 bp mRNA linear EST 02-DEC-2001  
 DEFINITION sg39a05.y1 Gm-c1025 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-c1025-681 5' similar to SW:P27084 SUPEROXIDE DISMUTASE [MN], MITOCHONDRIAL PRECURSOR, mRNA sequence.  
 ACCESSION BI942249  
 VERSION BI942249.1 GI:16277187  
 KEYWORDS EST.  
 SOURCE soybean.  
 ORGANISM Glycine max  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.  
 1 (bases 1 to 683)  
 Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Corvett, V., Rhanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Stepcoe, M., Theising, B., Allen, M., Bowers, Y., Peterson, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Rlter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann

TITLE R., Waterston, R. and Wilson, R.  
 JOURNAL Public Soybean EST Project  
 COMMENT Unpublished (1999)  
 Contact: Shoemaker R/Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available through: Resgen, Invitrogen Corp, 2130 South Memorial Parkway Huntsville, AL 35891 For further information call: (800)-533-4363 or contact via email: ctw@resgen.com  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 483.  
 FEATURES  
 source  
 1.683  
 /organism="Glycine max"  
 /db\_xref="taxon:3847"  
 /clone="GENOME SYSTEMS CLONE ID: Gm-c1025-681"  
 /clone.lib="Gm-c1025"  
 /tissue\_type="Hypocotyl, 3 day seedlings"  
 /lab\_host="DH10B"  
 /note="Vector: pT7T3Bac (pT7T3, Pharmacia); Site 1: EcoRI; Site 2: NotI; This cDNA library was constructed from mRNA isolated from hypocotyl tissue of 3 day old seedlings. Complementary DNA was synthesized from mRNA using a poly (dI) primer with a NotI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by digestion with EcoRI and NotI. The cDNA fragments were directionally cloned into the EcoRI-NotI restriction site of the pT7T3-Pac vector. The ligated cDNA fragments were transformed into DH10B host cells (Gibco BRL). This library was constructed by Dr. Randy Shoemaker."  
 BASE COUNT 174 a 183 c 175 g 151 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 0.00691 Length: 683  
 Score: 79.00 Matches: 13  
 Percent Similarity: 80.77% Conservative: 8  
 Best Local Similarity: 50.00% Mismatches: 5  
 Query Match: 54.48% Indels: 0  
 DB: 13 Gaps: 0  
 US-09-987-190-2 (1-30) x BI942249 (1-683)  
 QY 2 TyrSerleuProGluLeuAspTYRGIuPheSerAlaThrGluProTYRleSergIy 21  
 DB 150 TACACGCTCCCGATCTGATTAAGACTATGCGCTCTGAGCGCACCATCACGCGCGA 209  
 QY 22 IleAsnGluIle\*\*TYR 27  
 DB 210 ATCATGACGCTGCACAC 227  
 Search completed: April 8, 2003, 03:49:53  
 Job time : 283.532 secs







Query Match: 54.48% Indels: 0  
DB: 10 Gaps: 0  
US-09-987-190-2 (1-30) x US-09-878-574-2556 (1-360)  
QY 2 TySerLeuProGluLeuApyTyGluPheSerAlaThrGluProTyrIleSerGlyGln 21  
DB 150 TACACGCTCCCGCATCTGATTCGACTATGCGCTCTGGAGCCAGCATCAGCGCGGAA 209  
QY 22 IleAsnGluIle\*\*Tyr 27  
DB 210 ATCATGCACTGCACAC 227  
RESULT 2  
US-09-878-574-503  
; Sequence 503, Application US/09878574  
; Patent No. US2002010548A1  
; GENERAL INFORMATION:  
; APPLICANT: Byrum, Joseph R.  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Thompson, Michael D.  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(15401)B  
; CURRENT APPLICATION NUMBER: US/09/878,574  
; CURRENT FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 09/333,535  
; PRIOR FILING DATE: 1999-06-14  
; NUMBER OF SEQ ID NOS: 15775  
; SEQ ID NO 503  
; LENGTH: 390  
; TYPE: DNA  
; ORGANISM: Glycine max  
; OTHER INFORMATION: Clone ID: LIB3028-050-Q1-B1-C12  
US-09-878-574-503  
Alignment Scores:  
Pred. No.: 6.85e-06 Length: 390  
Score: 79.00 Matches: 13  
Percent Similarity: 80.77% Conservative: 8  
Best Local Similarity: 50.00% Mismatches: 5  
Query Match: 54.48% Indels: 0  
DB: 10 Gaps: 0  
US-09-987-190-2 (1-30) x US-09-878-574-503 (1-390)  
QY 2 TySerLeuProGluLeuApyTyGluPheSerAlaThrGluProTyrIleSerGlyGln 21  
DB 140 TACACGCTCCCGCATCTGATTCGACTATGCGCTCTGGAGCCAGCATCAGCGCGGAA 199  
QY 22 IleAsnGluIle\*\*Tyr 27  
DB 200 ATCATGCACTGCACAC 217  
RESULT 3  
US-09-738-626-3227  
; Sequence 3227, Application US/09738626  
; Publication No. US20020197605A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAGAWA, SATOSHI  
; APPLICANT: MIZOGUCHI, HIROSHI  
; APPLICANT: ANDO, SEIKO  
; APPLICANT: HAYASHI, MIKIRO  
; APPLICANT: OCHIAI, KEIHO  
; APPLICANT: YOKOI, HARUHIKO  
; APPLICANT: TATEISHI, NAKO  
; APPLICANT: SENOH, AKIHITO  
; APPLICANT: IKEDA, MASATO  
; APPLICANT: OZAKI, AKIO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-125  
; CURRENT APPLICATION NUMBER: US/09/738,626  
; CURRENT FILING DATE: 2000-12-18

; PRIOR APPLICATION NUMBER: JP 99/377484  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: JP 00/159162  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: JP 00/280988  
; PRIOR FILING DATE: 2000-08-03  
; NUMBER OF SEQ ID NOS: 7059  
; SOFTWARE: PatentIn ver. 3.0  
; SEQ ID NO 3227  
; LENGTH: 600  
; TYPE: DNA  
; ORGANISM: Corynebacterium glutamicum  
US-09-738-626-3227  
Alignment Scores:  
Pred. No.: 2.82e-05 Length: 600  
Score: 77.00 Matches: 13  
Percent Similarity: 74.07% Conservative: 7  
Best Local Similarity: 48.15% Mismatches: 7  
Query Match: 53.10% Indels: 0  
DB: 9 Gaps: 0  
US-09-987-190-2 (1-30) x US-09-738-626-3227 (1-600)  
QY 2 TySerLeuProGluLeuApyTyGluPheSerAlaThrGluProTyrIleSerGlyGln 21  
DB 10 TACGAAGCTCCAGAACTGACATGCGATGAGAGCGCTCTGAGCCACACATCCCGCTGAA 69  
QY 22 IleAsnGluIle\*\*TyrThr 28  
DB 70 ATCATGAGCTTCACCATCC 90  
RESULT 4  
US-09-818-564-1  
; Sequence 1, Application US/09818564  
; Patent No. US20020137151A1  
; GENERAL INFORMATION:  
; APPLICANT: MERKAM, MURIEL  
; APPLICANT: GUYONVARCH, ARMEI  
; APPLICANT: MARX, ACHIM  
; TITLE OF INVENTION: A PROCESS FOR THE FERMENTATIVE PREPARATION OF METABOLIC  
; TITLE OF INVENTION: PRODUCTS AND FOR THE NUCLEOTIDE SEQUENCES ENCODING FOR  
; TITLE OF INVENTION: THE SOD GENE  
; FILE REFERENCE: 21123/27841/MAS  
; CURRENT APPLICATION NUMBER: US/09/818,564  
; CURRENT FILING DATE: 2001-03-28  
; PRIOR APPLICATION NUMBER: 09/373,731  
; PRIOR FILING DATE: 1999-08-13  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 1143  
; TYPE: DNA  
; ORGANISM: Corynebacterium melassecola  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (338)..(937)  
; OTHER INFORMATION: ATCC 17965  
US-09-818-564-1  
Alignment Scores:  
Pred. No.: 6.51e-05 Length: 1143  
Score: 77.00 Matches: 13  
Percent Similarity: 74.07% Conservative: 7  
Best Local Similarity: 48.15% Mismatches: 7  
Query Match: 53.10% Indels: 0  
DB: 10 Gaps: 0  
US-09-987-190-2 (1-30) x US-09-818-564-1 (1-1143)  
QY 2 TySerLeuProGluLeuApyTyGluPheSerAlaThrGluProTyrIleSerGlyGln 21  
DB 347 TACGAAGCTCCAGAACTGACATGCGATGAGAGCGCTCTGAGCCACACATCCCGCTGAA 406

```

Oy      22 ileaangluile**TyrrThr 28
       |||::: ::::
Db      407 ATCATGAGACTTCACCACTCC 427

RESULT 5
US-09-960-352-5895
; Sequence 5895, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengping
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
FILE REFERENCE: 16511_006/37-21(10298)C
CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 15112
SEQ ID NO 5895
LENGTH: 224
TYPE: DNA
ORGANISM: Bos taurus
OTHER INFORMATION: Clone ID: 25-LIB34-057-Q1-E1-G1
US-09-960-352-5895

Alignment Scores:
Pred. No.:          2,85e-05           Length:         224
Score:              74.00             Matches:         12
Percent Similarity: 78.57%            Conservative:    10
Best Local Similarity: 42.86%          Mismatches:     6
Query Match:        51.03%            Indels:         0
DB:                 10                Gaps:           0

US-09-987-190-2 (1-30) x US-09-960-352-5895 (1-224)
Oy      1  LysTyrSerLeuProGluIleuAspTyrGluPheSerAlaThrgluProTyrIleSergly 20
       |||||:::|||||:::|||||:::|||||:::
Db      140 AAGCACAGCCTCCCCGACTTGCGTAACACTACGGCGCCCTGAGCCGCATCAACGCG 199

Oy      21 gniilaangluile**TyrrThr 28
       |||::: ::::
Db      200 CAGATCATGCAGCTGCACCACAGC 223

RESULT 6
US-09-960-352-3021
; Sequence 3021, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengping
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
FILE REFERENCE: 16511_006/37-21(10298)C
CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 15112
SEQ ID NO 3021
LENGTH: 259
TYPE: DNA
ORGANISM: Bos taurus
OTHER INFORMATION: Clone ID: 13-LIB34-076-Q1-E1-D1
US-09-960-352-3021

Alignment Scores:
Pred. No.:          3.44e-05           Length:         259
Score:              74.00             Matches:         12
Percent Similarity: 78.57%            Conservative:    10
Best Local Similarity: 42.86%          Mismatches:     6
Query Match:        51.03%            Indels:         0

```

```

US-09-987-190-2 (1-30) x US-09-960-352-3021 (1-259)
OY 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 142 AAGCAGACGCTCCCGACGCTCCGCTACGACTACGCGCGCCCTGAGACCATCATCAGCG 201
OY 21 GlnIleAsnGluIle***TyrThr 28
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 202 CAGATCATGCAGCTGCACACACAGC 225

RESULT 7
US-09-993-333-11
/ Sequence 11, Application US/09993333
/ Patent No. US20020156040A1
/ GENERAL INFORMATION:
/ APPLICANT: Oberley, Larry Wayne
/ APPLICANT: Oberley, Christine J.
/ APPLICANT: Smith, Benjamin Barnes
/ TITLE OF INVENTION: Reduction of antioxidant enzyme levels in tumor cells using antisense
/ FILE REFERENCE: 875,042US1
/ CURRENT APPLICATION NUMBER: US/09/993,333
/ CURRENT FILING DATE: 2001-11-14
/ PRIOR APPLICATION NUMBER: US 60/248,328
/ PRIOR FILING DATE: 2000-11-14
/ NUMBER OF SEQ ID NOS: 12
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 11
/ LENGTH: 325
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-993-333-11

Alignment Scores:
Pred. NO: 4.61e-05 Length: 325
Score: 74.00 Matches: 12
Percent Similarity: 78.57% Conservative: 10
Best Local Similarity: 42.86% Mismatches: 6
Query Match: 51.03% Indels: 0
DB: 9 Gaps: 0

US-09-987-190-2 (1-30) x US-09-993-333-11 (1-325)
OY 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 113 AAGCAGACGCTCCCGACGCTCCGCTACGACTACGCGCGCCCTGAGACCATCATCAGCG 172
OY 21 GlnIleAsnGluIle***TyrThr 28
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 173 CAGATCATGCAGCTGCACACACAGC 196

RESULT 8
US-09-960-352-14744
/ Sequence 14744, Application US/09960352
/ Patent No. US20020137139A1
/ GENERAL INFORMATION:
/ APPLICANT: Warren, Wesley C.
/ APPLICANT: Tao, Nengping
/ APPLICANT: Byatt, John C.
/ APPLICANT: Mathialagan, Nagappan
/ TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
/ TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
/ FILE REFERENCE: 16511,006/37-21(10298)C
/ CURRENT APPLICATION NUMBER: US/09/960,352
/ CURRENT FILING DATE: 2001-09-24
/ NUMBER OF SEQ ID NOS: 15112
/ SEQ ID NO 14744
/ LENGTH: 424
/ TYPE: DNA
/ ORGANISM: Bos taurus
/ OTHER INFORMATION: Clone ID: 63-LTB3057-022-Q1-K1-H4

```

US-09-960-352-14744

Alignment Scores:

Pred. No.:	6	Size-05	Length:	424
Score:	74.00	Matches:	12	
Percent Similarity:	78.57%	Conservative:	10	
Best Local Similarity:	42.86%	Mismatches:	6	
Query Match:	51.03%	Indels:	0	
DB:	10	Gaps:	0	

US-09-987-190-2 (1-30) x US-09-960-352-14744 (1-424)

QY 1 LyeTySerIeuProGluLeuApyTyGluPhseSerAlathrGluProTyrlleSergly 20

Db 121 AAGCAGACGCTCCCGACCTGCGACCTACGACGCGCCCTGAGCCGACATCAACGCG 180

QY 21 GlnlleAangluile\*\*TyThr 28

Db 181 CAGATCATGCGCTGCACACGACG 204

RESULT 9

US-10-044-090-836

Sequence 836, Application US/10044090

Patent No. US20020137081A1

GENERAL INFORMATION:

APPLICANT: Olga Bandman

TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION

FILE REFERENCE: PA-0028 US

CURRENT APPLICATION NUMBER: US/10/044,090

CURRENT FILING DATE: 2002-01-09

NUMBER OF SEQ ID NOS: 850

SOFTWARE: PERL Program

SEQ ID NO 836

LENGTH: 972

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

OTHER INFORMATION: Incyte ID No. US20020137081A1 010001CB1

US-10-044-090-836

Alignment Scores:

Pred. No.:	0.000191	Length:	972
Score:	74.00	Matches:	12
Percent Similarity:	78.57%	Conservative:	10
Best Local Similarity:	42.86%	Mismatches:	6
Query Match:	51.03%	Indels:	0
DB:	12	Gaps:	0

US-09-987-190-2 (1-30) x US-10-044-090-836 (1-972)

QY 1 LyeTySerIeuProGluLeuApyTyGluPhseSerAlathrGluProTyrlleSergly 20

Db 166 AAGCAGACGCTCCCGACCTGCGACCTACGACGCGCCCTGAGCCGACATCAACGCG 225

QY 21 GlnlleAangluile\*\*TyThr 28

Db 226 CAGATCATGCGCTGCACACGACG 249

RESULT 10

US-09-954-456-1822

Sequence 1822, Application US/09954456

Patent No. US20020115057A1

GENERAL INFORMATION:

APPLICANT: Young, Paul

TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand

FILE REFERENCE: 689290-76

CURRENT APPLICATION NUMBER: US/09/954,456

CURRENT FILING DATE: 2001-09-18

PRIOR APPLICATION NUMBER: US/60/233,617

PRIOR FILING DATE: 2000-09-18

PRIOR APPLICATION NUMBER: US/60/234,052

PRIOR FILING DATE: 2000-09-20

PRIOR APPLICATION NUMBER: US/60/234,923

PRIOR FILING DATE: 2000-09-25

PRIOR APPLICATION NUMBER: US/60/235,134

PRIOR FILING DATE: 2000-09-25

PRIOR APPLICATION NUMBER: US/60/235,637

PRIOR FILING DATE: 2000-09-26

PRIOR APPLICATION NUMBER: US/60/235,638

PRIOR FILING DATE: 2000-09-26

PRIOR APPLICATION NUMBER: US/60/235,711

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: US/60/235,720

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: US/60/235,840

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: US/60/235,863

PRIOR FILING DATE: 2000-09-27

NUMBER OF SEQ ID NOS: 2276

SOFTWARE: PatentIn version 3.0

SEQ ID NO 1822

LENGTH: 1026

TYPE: DNA

ORGANISM: Homo sapiens

US-09-954-456-1822

Alignment Scores:

Pred. No.:	0.000205	Length:	1026
Score:	74.00	Matches:	12
Percent Similarity:	78.57%	Conservative:	10
Best Local Similarity:	42.86%	Mismatches:	6
Query Match:	51.03%	Indels:	0
DB:	10	Gaps:	0

US-09-987-190-2 (1-30) x US-09-954-456-1822 (1-1026)

QY 1 LyeTySerIeuProGluLeuApyTyGluPhseSerAlathrGluProTyrlleSergly 20

Db 77 AAGCAGACGCTCCCGACCTGCGACCTACGACGCGCCCTGAGCCGACATCAACGCG 136

QY 21 GlnlleAangluile\*\*TyThr 28

Db 137 CAGATCATGCGCTGCACACGACG 160

RESULT 11

US-09-917-800A-1627

Sequence 1627, Application US/09917800A

Patent No. US20020119462A1

GENERAL INFORMATION:

APPLICANT: Mendrick, Donna

APPLICANT: Porter, Mark

APPLICANT: Johnson, Kory

APPLICANT: Castle, Arthur

APPLICANT: Elashoff, Michael

APPLICANT: Gene Logic, Inc.

TITLE OF INVENTION: Molecular Toxicology Modeling

FILE REFERENCE: 44921-5038-US

CURRENT APPLICATION NUMBER: US/09/917,800A

CURRENT FILING DATE: 2001-07-31

PRIOR APPLICATION NUMBER: US 60/222,040

PRIOR FILING DATE: 2000-07-31

PRIOR APPLICATION NUMBER: US 60/222,880

PRIOR FILING DATE: 2000-11-02

PRIOR APPLICATION NUMBER: US 60/290,029

PRIOR FILING DATE: 2001-05-11

PRIOR APPLICATION NUMBER: US 60/290,645

PRIOR FILING DATE: 2001-05-15

PRIOR APPLICATION NUMBER: US 60/292,336

PRIOR FILING DATE: 2001-05-22

PRIOR APPLICATION NUMBER: US 60/295,798

PRIOR FILING DATE: 2001-06-06

PRIOR APPLICATION NUMBER: US 60/297,457

PRIOR FILING DATE: 2001-06-13

PRIOR APPLICATION NUMBER: US 60/298,884

```

; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1627
; LENGTH: 1492
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM_017051
US-09-917-800A-1627

Alignment Scores:
Pred. No.: 0.000332 Length: 1492
Score: 74.00 Matches: 12
Percent Similarity: 78.57% Conservative: 10
Best Local Similarity: 42.86% Mismatches: 6
Query Match: 51.03% Indels: 0
Gaps: 0

US-09-987-190-2 (1-30) x US-09-917-800A-1627 (1-1492)

QY 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 141 MGGCAGAGCTCGCTTACGACTATGCGCGCTGAGCGGCGGACATTAACGCG 200

QY 21 GlnIleAsnGluIle**TyrThr 28
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 201 CAGATCATGCAGCTGCACGACGAC 224

RESULT 12
US-09-987-190-8
; Sequence 8, Application US/09987190
; Patent No. US20020058293A1
; GENERAL INFORMATION:
; APPLICANT: TAKESAKO, Kazutoh
; APPLICANT: MIZUTANI, Shigetoshi
; APPLICANT: ENDO, Masahiro
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: FUNGAL ANTIGENS AND PROCESS FOR PRODUCING THE SAME
; FILE REFERENCE: 1432-0502P
; CURRENT APPLICATION NUMBER: US/09/987,190
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: 09/262,856
; PRIOR FILING DATE: 1999-03-04
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 721
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-987-190-8

Alignment Scores:
Pred. No.: 0.000469 Length: 721
Score: 71.00 Matches: 13
Percent Similarity: 93.75% Conservative: 2
Best Local Similarity: 81.25% Mismatches: 1
Query Match: 48.97% Indels: 0
Gaps: 0

US-09-987-190-2 (1-30) x US-09-987-190-8 (1-721)

QY 13 AlaThrGluProTyrIleSerGlyGlnIleAsnGluIle**TyrThr 28
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 GGCACGAGACCGTACATCACAGACAAAGACGAATTCACCTACTACT 48

RESULT 13
US-09-770-445-745
; Sequence 745, Application US/09770445
; Patent No. US2002023281A1
; GENERAL INFORMATION:
; APPLICANT: Gorlach, Jörn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameake, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Krickler, Maja
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hurlban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; FILE REFERENCE: 2023US (PARA-012PRV)
; CURRENT APPLICATION NUMBER: US/09/770,445
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 60/178,472
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 745
; LENGTH: 817
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-770-445-745

Alignment Scores:
Pred. No.: 0.000846 Length: 817
Score: 70.00 Matches: 12
Percent Similarity: 76.92% Conservative: 8
Best Local Similarity: 46.15% Mismatches: 6
Query Match: 48.28% Indels: 0
Gaps: 0

US-09-987-190-2 (1-30) x US-09-770-445-745 (1-817)

QY 2 TyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGlyGln 21
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 48 TTACGCTTCCTGATCTTCCTTACGACTATGCGCGGCGGACGCGCATTAAGTGAAG 107

QY 22 IleAsnGluIle**Tyr 27
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 108 ATCATGCAGATTCATCAC 125

RESULT 14
US-09-727-855B-6
; Sequence 6, Application US/09727855B
; Patent No. US20020168703A1
; GENERAL INFORMATION:
; APPLICANT: HOSHINO, Tatsuo
; APPLICANT: OJIMA, Kazuyuki
; APPLICANT: SETOGUCHI, Yutaka
; TITLE OF INVENTION: PROCESS FOR THE MANUFACTURE OF CAROTENOIDS AND BIOLOGICALLY USEFUL
; FILE REFERENCE: C38435/111694
; CURRENT APPLICATION NUMBER: US/09/727,855B
; CURRENT FILING DATE: 2000-12-01
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 597
; TYPE: DNA
; ORGANISM: Phaffia rhodozyma
; FEATURE:
; NAME/KEY: CDS

```

LOCATION: (1)..(594)  
OTHER INFORMATION:  
US-09-727-855B-6

Alignment Scores:  
Pred. No.: 0.00204  
Score: 67.00  
Percent Similarity: 76.19%  
Best Local Similarity: 57.14%  
Query Match: 46.21%  
DB: 9  
Length: 597  
Matches: 12  
Conservative: 4  
Mismatch: 5  
Indels: 0  
Gaps: 0

US-09-987-190-2 (1-30) x US-09-727-855B-6 (1-597)

QY 2 TySerleupProgluLeuAapTyGluPheserAlaThrgluProTyrlleSerglyGln 21  
DB 10 TACACTCTCCCGACCTTCCCTTACGCTTACGATGCTTGAGCTTACATCTTAAGAA 69

QY 22 1le 22  
DB 70 ATC 72

RESULT 15  
US-09-727-855B-4  
Sequence 4, Application US/09727855B  
Patent No. US20020168703A1  
GENERAL INFORMATION:  
APPLICANT: HOSHINO, Kazuo  
APPLICANT: OJIMA, Kazuyuki  
APPLICANT: SETOGUCHI, Yutaka  
TITLE OF INVENTION: PROCESS FOR THE MANUFACTURE OF CAROTENOIDS AND BIOLOGICALLY USEFUL  
FILE REFERENCE: C38435/111694  
CURRENT APPLICATION NUMBER: US/09/727,855B  
CURRENT FILING DATE: 2000-12-01  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 4  
LENGTH: 669  
TYPE: DNA  
ORGANISM: Phaeofila rhodozyma  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(666)  
OTHER INFORMATION: n or X = A, C, G or T

US-09-727-855B-4

Alignment Scores:  
Pred. No.: 0.00363  
Score: 66.00  
Percent Similarity: 67.86%  
Best Local Similarity: 46.43%  
Query Match: 45.52%  
DB: 9  
Length: 669  
Matches: 13  
Conservative: 6  
Mismatch: 9  
Indels: 0  
Gaps: 0

US-09-987-190-2 (1-30) x US-09-727-855B-4 (1-669)

QY 1 TyTySerleupProgluLeuAapTyGluPheserAlaThrgluProTyrlleSerglyGln 20  
DB 76 AAGCATACCTCGCTGAGCTTCCCTTACGCTTACGATGCTTGAGCTTCCATCTCAAG 135

QY 21 GlnleAangluile\*\*TyThr 28  
DB 136 GAGTATGACCTTACACACACC 159

RESULT 16  
US-09-974-300-2052  
Sequence 2052, Application US/09974300  
Patent No. US20020146721A1  
GENERAL INFORMATION:  
APPLICANT: Berka, Randy M.  
APPLICANT: Clausen, Ib Groth  
TITLE OF INVENTION: Methods For Monitoring Multiple Gene

TITLE OF INVENTION: Expression  
FILE REFERENCE: 10085.500-US  
CURRENT APPLICATION NUMBER: US/09/974,300  
CURRENT FILING DATE: 2001-10-05  
PRIOR APPLICATION NUMBER: 09/680,598  
PRIOR FILING DATE: 2000-10-06  
PRIOR APPLICATION NUMBER: 60/279,526  
PRIOR FILING DATE: 2001-03-27  
NUMBER OF SEQ ID NOS: 8481  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2052  
LENGTH: 838  
TYPE: DNA  
ORGANISM: Bacillus licheniformis

US-09-974-300-2052

Alignment Scores:  
Pred. No.: 0.00485  
Score: 66.00  
Percent Similarity: 66.67%  
Best Local Similarity: 48.15%  
Query Match: 45.52%  
DB: 10  
Length: 838  
Matches: 13  
Conservative: 5  
Mismatch: 9  
Indels: 0  
Gaps: 0

US-09-987-190-2 (1-30) x US-09-974-300-2052 (1-838)

QY 2 TySerleupProgluLeuAapTyGluPheserAlaThrgluProTyrlleSerglyGln 21  
DB 238 CATCGCTCGCAGCTCCCTTACTGATTCAGCTTGAGCTTATTTAAAGAA 297

QY 22 1leAangluile\*\*TyThr 28  
DB 298 ATTATGTTCTTCATCATACA 318

RESULT 17  
US-09-070-927A-425  
Sequence 425, Application US/09070927A  
Patent No. US20020120116A1  
GENERAL INFORMATION:  
APPLICANT: Charles A. Kunsch  
Patrick J. Dillon  
Steven Barash  
TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides  
NUMBER OF SEQUENCES: 982  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/070,927A  
FILING DATE: 04-May-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/046,655  
FILING DATE: 1997-05-16  
APPLICATION NUMBER: 60/044,031  
FILING DATE: 1997-05-06  
APPLICATION NUMBER: 60/066,009  
FILING DATE: 1997-11-14  
ATTORNEY/AGENT INFORMATION:  
NAME: Kenley K. Hoover  
REGISTRATION NUMBER: 40,302  
REFERENCE/DOCKET NUMBER: PB369  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504

```

;
; TELEFAX: (301) 309-8512
;
; INFORMATION FOR SEQ ID NO: 425:
;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 767 base pairs
;
; TYPE: nucleic acid
;
; STRANDEDNESS: double
;
; TOPOLOGY: linear
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 425:
US-09-070-927A-425

Alignment Scores:
Pred. No.: 0.024 Length: 767
Score: 62.00 Matches: 11
Percent Similarity: 76.47% Conservative: 2
Best Local Similarity: 64.71% Mismatches: 4
Query Match: 42.76% Indels: 0
DB: 10 Gaps: 0

09-987-190-2 (1-30) x US-09-070-927A-425 (1-767)

Qy 2 TySeriesProgluenuapTyrgluPheserAlathrgluProtyrIleserGlyIn 18
124 TAGACATTCACAGATTACCATGCTTATGATGACATTAGACCTTACATT 174

RESULT 18
US-09-790-988-1
; Sequence 1, Application US/09790988
; Patent No. US20020127687A1
; GENERAL INFORMATION:
; APPLICANT: SHIGEMOBU, SHUJI
; APPLICANT: MATANABE, HIDEKI
; APPLICANT: HATTORI, MASAHIRA
; APPLICANT: SAKAKI, YOSHIYUKI
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
; FILE REFERENCE: 081356/0159
; CURRENT APPLICATION NUMBER: US/09/790,988
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: JP2000-107160
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 640681
; TYPE: DNA
; ORGANISM: Buchnera sp.
; US-09-790-988-1

Alignment Scores:
Pred. No.: 1.93e+03 Length: 640681
Score: 56.00 Matches: 10
Percent Similarity: 59.26% Conservative: 6
Best Local Similarity: 37.04% Mismatches: 11
Query Match: 38.62% Indels: 0
DB: 10 Gaps: 0

US-09-987-190-2 (1-30) x US-09-790-988-1 (1-640681)

Qy 2 TySeriesProgluenuapTyrgluPheserAlathrgluProtyrIleserGlyIn 21
DB 204469 TAGGTTCTTCCTTACCTTATCAATGCAATGCGTTAGAACATTTTTGTGTAAGAA 204528

Qy 22 Ileaungluile**TyThr 28
DB 204529 ACTATGAAATTCATCTACT 204549

RESULT 19
US-09-770-445-976
; Sequence 976, Application US/09770445
; Patent No. US20020023281A1
; GENERAL INFORMATION:
; APPLICANT: Gorlach, Jörn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; US-09-770-445-976
```

```

; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Mathew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Kricker, Maja
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hurban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; FILE REFERENCE: 2023US (PAPA-012PRV)
; CURRENT APPLICATION NUMBER: US/09/770,445
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 60/178,472
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 976
; LENGTH: 757
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(757)
; OTHER INFORMATION: n = A,T,C or G
US-09-770-445-976

Alignment Scores:
Pred. No.: 0.729 Length: 757
Score: 54.00 Matches: 10
Percent Similarity: 75.00% Conservative: 5
Best Local Similarity: 50.00% Mismatches: 5
Query Match: 37.24% Indels: 0
DB: 10 Gaps: 0

US-09-987-190-2 (1-30) x US-09-770-445-976 (1-757)

Qy 1 TySeriesProgluenuapTyrgluPheserAlathrgluProtyrIleserGly 20
DB 236 AAAGAAGCATTTAGAGAGCTTCCACACAAATTCACATCACGACCGTTATTTCGGGT 295

RESULT 20
US-08-781-986A-777/C
; Sequence 777, Application US/08781986A
; Publication No. US20030054436A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
```





```

; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: US/09/565,391
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: PCT/US99/26409
; PRIOR FILING DATE: 1999-11-09
; PRIOR APPLICATION NUMBER: 60/108,207
; PRIOR FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 115
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
; LENGTH: 1044
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-948-820-29

Alignment Scores:
Pred. No.: 1 7
Score: 53.00 Length: 1044
Percent Similarity: 58.33% Matches: 11
Best Local Similarity: 45.83% Conservative: 3
Query Match: 36.55% Mismatches: 8
DB: Gaps: 2
Indels: 1

US-09-987-190-2 (1-30) x US-09-948-820-29 (1-1044)

QY 1 TySerLeuProGluLeuAspTyr-----GluPheSerAlaThrGluProTyrIle 18
Db 642 AAADACTGCTACTGATTTGACCACTCCTCATTCATTCACCTCCACCTCCTACACCATATATT 701

QY 19 SerGlyGlnIle 22
Db 702 TCAGGTCTCTT 713

RESULT 24
US-09-974-300-6347
; Sequence 6347, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6347
; LENGTH: 385
; TYPE: DNA
; ORGANISM: Bacillus clausii
US-09-974-300-6347

Alignment Scores:
Pred. No.: 0.715 Length: 385
Score: 52.00 Matches: 10
Percent Similarity: 70.59% Conservative: 2
Best Local Similarity: 58.82% Mismatches: 5
Query Match: 35.86% Indels: 0
DB: Gaps: 0

US-09-987-190-2 (1-30) x US-09-974-300-6347 (1-385)

QY 2 TySerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIle 18
Db 7 TACAACTACCAAGATTGCTTAGCGGCAATGCACTTGACCGCATATT 57

RESULT 25
US-09-822-849A-208
```

```

; Sequence 208, Application US/09822849A
; Patent No. US20020045170A1
; GENERAL INFORMATION:
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalakara
; APPLICANT: Graham, James R.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6403
; CURRENT APPLICATION NUMBER: US/09/822,849A
; CURRENT FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: 60/195,582
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 598
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 208
; LENGTH: 2454
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-849A-208

Alignment Scores:
Pred. No.: 12.1 Length: 2454
Score: 51.00 Matches: 10
Percent Similarity: 60.00% Conservative: 2
Best Local Similarity: 50.00% Mismatches: 8
Query Match: 35.17% Indels: 0
DB: Gaps: 0

US-09-987-190-2 (1-30) x US-09-822-849A-208 (1-2454)

QY 1 TySerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20
Db 1797 AGATATAGCTCCAGAGATTCTCTATATGCAATCGTGAAAGAACCATATCCAGCAGGA 1856

RESULT 26
US-10-091-504-2314
; Sequence 2314, Application US/10091504
; Publication No. US20030059908A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007C1
; CURRENT APPLICATION NUMBER: US/10/091,504
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2442
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2314
; LENGTH: 32249
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-091-504-2314

Alignment Scores:
Pred. No.: 341 Length: 32249
Score: 51.00 Matches: 9
Percent Similarity: 68.18% Conservative: 6
Best Local Similarity: 40.91% Mismatches: 7
Query Match: 35.17% Indels: 0
DB: Gaps: 0

US-09-987-190-2 (1-30) x US-10-091-504-2314 (1-32249)

QY 2 TySerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGlyGln 21
Db 31767 TATAGCTCCACATCATTCATTATATATATGAGAGCCCTTTATCTCTCCAA 31826
```

QY 22 ILeaen 23  
Db 31827 GTAGC 31832

RESULT 27  
US-09-764-869-2314  
Sequence 2314, Application US/09764869  
Patent No. US20020061521A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PC007  
CURRENT APPLICATION NUMBER: US/09/764,869  
CURRENT FILING DATE: 2001-01-17  
Prior application data removed - refer to PALM or file wrapper  
NUMBER OF SEQ ID NOS: 2442  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 2314  
LENGTH: 32249  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-764-869-2314 (1-32249)

Alignment Scores:  
Pred. No.: 341 Length: 32249  
Score: 51.00 Matches: 9  
Percent Similarity: 68.18% Conservative: 6  
Best Local Similarity: 40.91% Mismatches: 7  
Query Match: 35.17% Indels: 0  
Gaps: 0

QY 2 TySerLeuProGluLeuAAsPyrGluPheSerAlaThrGluProTyrIleSerGlyln 21  
Db 31767 TATAGCCTCCACATCATACATTTTCAATTAATATAGAGAGCCCTTTATCTCTCCAA 31826

QY 22 ILeaen 23  
Db 31827 GTAGC 31832

RESULT 28  
US-10-012-896-851  
Sequence 851, Application US/10012896  
Publication No. US20020183251A1  
GENERAL INFORMATION:  
APPLICANT: Xu, Jiangchun  
APPLICANT: Dillon, Davin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Jiang, Yuguu  
APPLICANT: Kalos, Michael D.  
APPLICANT: Retter, Marc W.  
APPLICANT: Stolk, John A.  
APPLICANT: Day, Craig H.  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Carter, Darrick  
APPLICANT: Li, Samuel X.  
APPLICANT: Wang, Aijun  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Hepler, William T.  
APPLICANT: Henderson, Robert A.  
APPLICANT: Hural, John  
APPLICANT: McNeill, Patricia D.  
APPLICANT: Houghton, Raymond L.  
APPLICANT: Vinals de Bassols, Carlota  
APPLICANT: Foy, Teresa  
APPLICANT: Fanger, Gary R.  
APPLICANT: Mantanabe, Yoshihiro  
APPLICANT: Mesgher, Madeline Joy  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
FILE REFERENCE: 210121.427C27

US-09-987-190-2 (1-30) x US-10-012-896-851 (1-1203)

QY 2 TySerLeuProGluLeuAAsPyrGluPheSerAlaThrGluProTyrIleSerGlyln 21  
Db 832 TACCTCCTGCTGCATTCATGCGACACCATGCTGCGCCCTTACTGAGCACCAG 891

RESULT 29  
US-09-895-793-851  
Sequence 851, Application US/09895793  
Publication No. US20020192763A1  
GENERAL INFORMATION:  
APPLICANT: Xu, Jiangchun  
APPLICANT: Dillon, Davin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Jiang, Yuguu  
APPLICANT: Kalos, Michael D.  
APPLICANT: Retter, Marc W.  
APPLICANT: Stolk, John A.  
APPLICANT: Day, Craig H.  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Carter, Darrick  
APPLICANT: Li, Samuel X.  
APPLICANT: Wang, Aijun  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Hepler, William T.  
APPLICANT: Henderson, Robert A.  
APPLICANT: Hural, John  
APPLICANT: McNeill, Patricia D.  
APPLICANT: Houghton, Raymond L.  
APPLICANT: Vinals de Bassols, Carlota  
APPLICANT: Foy, Teresa  
APPLICANT: Fanger, Gary R.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
FILE REFERENCE: 210121.534C2  
CURRENT APPLICATION NUMBER: US/09/895,793  
CURRENT FILING DATE: 2001-06-29  
NUMBER OF SEQ ID NOS: 982  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 851  
LENGTH: 1203  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-895-793-851

Alignment Scores:  
Pred. No.: 7.38 Length: 1203  
Score: 50.00 Matches: 9  
Percent Similarity: 65.00% Conservative: 4  
Best Local Similarity: 45.00% Mismatches: 7  
Query Match: 34.48% Indels: 0  
Gaps: 0

US-09-987-190-2 (1-30) x US-09-895-793-851 (1-1203)

QY 2 TySerLeuProGluLeuAAsPyrGluPheSerAlaThrGluProTyrIleSerGlyln 21  
Db 832 TACCTCCTGCTGCATTCATGCGACACCATGCTGCGCCCTTACTGAGCACCAG 891

RESULT 29  
US-09-895-793-851  
Sequence 851, Application US/09895793  
Publication No. US20020192763A1  
GENERAL INFORMATION:  
APPLICANT: Xu, Jiangchun  
APPLICANT: Dillon, Davin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Jiang, Yuguu  
APPLICANT: Kalos, Michael D.  
APPLICANT: Retter, Marc W.  
APPLICANT: Stolk, John A.  
APPLICANT: Day, Craig H.  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Carter, Darrick  
APPLICANT: Li, Samuel X.  
APPLICANT: Wang, Aijun  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Hepler, William T.  
APPLICANT: Henderson, Robert A.  
APPLICANT: Hural, John  
APPLICANT: McNeill, Patricia D.  
APPLICANT: Houghton, Raymond L.  
APPLICANT: Vinals de Bassols, Carlota  
APPLICANT: Foy, Teresa  
APPLICANT: Fanger, Gary R.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
FILE REFERENCE: 210121.534C2  
CURRENT APPLICATION NUMBER: US/09/895,793  
CURRENT FILING DATE: 2001-06-29  
NUMBER OF SEQ ID NOS: 982  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 851  
LENGTH: 1203  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-895-793-851

Alignment Scores:  
Pred. No.: 7.38 Length: 1203  
Score: 50.00 Matches: 9  
Percent Similarity: 65.00% Conservative: 4  
Best Local Similarity: 45.00% Mismatches: 7  
Query Match: 34.48% Indels: 0  
Gaps: 0

US-09-987-190-2 (1-30) x US-09-895-793-851 (1-1203)

```
Oy 2 TySerLeupProGluLeuApyTyrgLuphSeRaIaThrgLupProTyrlleSerglygn 21
Db 832 TACCTCGTGGCTGCATTGACTGGACACCAAGTGGCCCTGACCTGAGGACCCAG 891

RESULT 30
US-09-895-814-851
; Sequence 851, Application US/09895814
; Publication No. US20020193296A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C26
; CURRENT APPLICATION NUMBER: US/09/895,814
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 851
; LENGTH: 1203
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-895-814-851

Alignment Scores:
Pred. No.: 7.38 Length: 1203
Score: 50.00 Matches: 9
Percent Similarity: 65.00% Conservative: 4
Best Local Similarity: 45.00% Mismatches: 7
Query Match: 34.48% Indels: 0
DB: 9 Gaps: 0

US-09-987-190-2 (1-30) x US-09-895-814-851 (1-1203)
Oy 2 TySerLeupProGluLeuApyTyrgLuphSeRaIaThrgLupProTyrlleSerglygn 21
Db 832 TACCTCGTGGCTGCATTGACTGGACACCAAGTGGCCCTGACCTGAGGACCCAG 891

RESULT 31
US-09-759-143-851
; Sequence 851, Application US/09759143
; Patent No. US20020022248A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
```

```
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C23
; CURRENT APPLICATION NUMBER: US/09/759,143
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 851
; LENGTH: 1203
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-759-143-851

Alignment Scores:
Pred. No.: 7.38 Length: 1203
Score: 50.00 Matches: 9
Percent Similarity: 65.00% Conservative: 4
Best Local Similarity: 45.00% Mismatches: 7
Query Match: 34.48% Indels: 0
DB: 10 Gaps: 0

US-09-987-190-2 (1-30) x US-09-759-143-851 (1-1203)
Oy 2 TySerLeupProGluLeuApyTyrgLuphSeRaIaThrgLupProTyrlleSerglygn 21
Db 832 TACCTCGTGGCTGCATTGACTGGACACCAAGTGGCCCTGACCTGAGGACCCAG 891

RESULT 32
US-09-780-669-851
; Sequence 851, Application US/09780669
; Patent No. US20020051977A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C24
; CURRENT APPLICATION NUMBER: US/09/780,669
; NUMBER OF SEQ ID NOS: 943
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 851
; LENGTH: 1203
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-780-669-851
```

```
Alignment Scores:
Pred. No.: 7.38 Length: 1203
Score: 50.00 Matches: 9
Percent Similarity: 65.00% Conservative: 4
Best Local Similarity: 45.00% Mismatches: 7
Query Match: 34.48% Indels: 0
DB: 10 Gaps: 0

US-09-987-190-2 (1-30) x US-09-780-669-851 (1-1203)

QY 2 TySerleupProGluLeuAspTyRgluPhseSerAlaThrGluProTyrIleSerGlyGln 21
DB 832 TACTCTGCTGCTGCATTGACGTGGACACCACTGCTGCGCCCTTACTGGGACCCAG 891

RESULT 33
US-09-822-827-851
; Sequence 851, Application US/09822827
; Patent No. US2002008180A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.534C1
; CURRENT APPLICATION NUMBER: US/09/822,827
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 851
; LENGTH: 1203
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-827-851

Alignment Scores:
Pred. No.: 7.38 Length: 1203
Score: 50.00 Matches: 9
Percent Similarity: 65.00% Conservative: 4
Best Local Similarity: 45.00% Mismatches: 7
Query Match: 34.48% Indels: 0
DB: 10 Gaps: 0

US-09-987-190-2 (1-30) x US-09-822-827-851 (1-1203)

QY 2 TySerleupProGluLeuAspTyRgluPhseSerAlaThrGluProTyrIleSerGlyGln 21
DB 832 TACTCTGCTGCTGCATTGACGTGGACACCACTGCTGCGCCCTTACTGGGACCCAG 891

RESULT 34
US-10-012-896-703
Sequence 703, Application US/10012896
Publication No. US20020183251A1
GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aljun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Huxal, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
```

```
; APPLICANT: Fanger, Gary R.
; APPLICANT: Wantanabe, Yoshihiro
; APPLICANT: Meagher, Madeline Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C27
; CURRENT APPLICATION NUMBER: US/10/012,896
; NUMBER OF SEQ ID NOS: 1011
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 703
; LENGTH: 2904
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-012-896-703

Alignment Scores:
Pred. No.: 23.1 Length: 2904
Score: 50.00 Matches: 9
Percent Similarity: 65.00% Conservative: 4
Best Local Similarity: 45.00% Mismatches: 7
Query Match: 34.48% Indels: 0
DB: 9 Gaps: 0

US-09-987-190-2 (1-30) x US-10-012-896-703 (1-2904)

QY 2 TySerleupProGluLeuAspTyRgluPhseSerAlaThrGluProTyrIleSerGlyGln 21
DB 40 TACTCTGCTGCTGCATTGACGTGGACACCACTGCTGCGCCCTTACTGGGACCCAG 99

RESULT 35
US-09-895-793-703
Sequence 703, Application US/09895793
Publication No. US20020192763A1
GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aljun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.534C2
; CURRENT APPLICATION NUMBER: US/09/895,793
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 703
; LENGTH: 2904
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-895-793-703

Alignment Scores:
Pred. No.: 23.1 Length: 2904
Score: 50.00 Matches: 9
```

```
Percent Similarity: 65.00% Conservative: 4
Best Local Similarity: 45.00% Mismatches: 7
Query Match: 34.48% Indels: 0
DB: 9 Gaps: 0

US-09-987-190-2 (1-30) x US-09-895-793-703 (1-2904)

QY 2 TySerLeuProGluLeuApyTYrGluPheSerAlaThrGluProTYrIleSerGlyGln 21
   |||||  ::|||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 40 TACCTCTGCTGCATYGAATGAGACACCAAGTGCCTGCGCCCTTACTGGGACACCCAG 99

RESULT 36
US-09-895-814-703
; Sequence 703, Application US/09895814
; Publication No. US20020193296A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darriick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carloca
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; APPLICANT: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C26
; CURRENT APPLICATION NUMBER: US/09/895,814
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 703
; LENGTH: 2904
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-895-814-703

Alignment Scores:
Pred. No.: 23.1 Length: 2904
Score: 50.00 Matches: 9
Percent Similarity: 65.00% Conservative: 4
Best Local Similarity: 45.00% Mismatches: 7
Query Match: 34.48% Indels: 0
DB: 9 Gaps: 0

US-09-987-190-2 (1-30) x US-09-895-814-703 (1-2904)

QY 2 TySerLeuProGluLeuApyTYrGluPheSerAlaThrGluProTYrIleSerGlyGln 21
   |||||  ::|||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 40 TACCTCTGCTGCATYGAATGAGACACCAAGTGCCTGCGCCCTTACTGGGACACCCAG 99

RESULT 37
US-09-759-143-703
; Sequence 703, Application US/09759143
; Patent No. US2002002248A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
```

```
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darriick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; APPLICANT: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C23
; CURRENT APPLICATION NUMBER: US/09/759,143
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 703
; LENGTH: 2904
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-759-143-703

Alignment Scores:
Pred. No.: 23.1 Length: 2904
Score: 50.00 Matches: 9
Percent Similarity: 65.00% Conservative: 4
Best Local Similarity: 45.00% Mismatches: 7
Query Match: 34.48% Indels: 0
DB: 10 Gaps: 0

US-09-987-190-2 (1-30) x US-09-759-143-703 (1-2904)

QY 2 TySerLeuProGluLeuApyTYrGluPheSerAlaThrGluProTYrIleSerGlyGln 21
   |||||  ::|||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 40 TACCTCTGCTGCATYGAATGAGACACCAAGTGCCTGCGCCCTTACTGGGACACCCAG 99

RESULT 38
US-09-780-669-703
; Sequence 703, Application US/09780669
; Patent No. US20020051977A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darriick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C24
; CURRENT APPLICATION NUMBER: US/09/780,669
; NUMBER OF SEQ ID NOS: 943
; SOFTWARE: FastSeq for Windows Version 3.0
```

SEQ ID NO 703  
LENGTH: 2904  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-780-669-703

Alignment Scores:  
Pred. No.: 23.1 Length: 2904  
Score: 50.00 Matches: 9  
Percent Similarity: 65.00% Conservative: 4  
Best Local Similarity: 45.00% Mismatches: 7  
Query Match: 34.48% Indels: 0  
DB: 10 Gaps: 0

US-09-987-190-2 (1-30) x US-09-780-669-703 (1-2904)

Qy 2 TySerleupProgluLeuApyTyrgluPheserAlaThrgluProTyrlleserGlyln 21  
Db 40 TACTCTGCTGCTGCATTGACGTGGACACCACTGCTGCCCTTACTGGGACCCAG 99

RESULT 39  
Sequence 703, Application US/09822827  
Patent No. US200208180A1  
GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
FILE REFERENCE: 210121.534C1  
CURRENT APPLICATION NUMBER: US/09/822,827  
NUMBER OF SEQ ID NOS: 982  
SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 703  
LENGTH: 2904  
TYPE: DNA

ORGANISM: Homo sapiens  
US-09-822-827-703

Alignment Scores:  
Pred. No.: 23.1 Length: 2904  
Score: 50.00 Matches: 9  
Percent Similarity: 65.00% Conservative: 4  
Best Local Similarity: 45.00% Mismatches: 7  
Query Match: 34.48% Indels: 0  
DB: 10 Gaps: 0

US-09-987-190-2 (1-30) x US-09-822-827-703 (1-2904)

Qy 2 TySerleupProgluLeuApyTyrgluPheserAlaThrgluProTyrlleserGlyln 21  
Db 40 TACTCTGCTGCTGCATTGACGTGGACACCACTGCTGCCCTTACTGGGACCCAG 99

RESULT 40  
US-09-838-785-1

Sequence 1, Application US/09838785  
Patent No. US2002009455A1  
GENERAL INFORMATION:

APPLICANT: Lau, Ted  
APPLICANT: Lin, Rick  
APPLICANT: Parkes, Debbie  
APPLICANT: Parry, Gordon  
APPLICANT: Schneider, Douglas  
APPLICANT: Steinhilber, Renate  
APPLICANT: Van Heule, Pam T  
APPLICANT: Wu, John

TITLE OF INVENTION: DNA Encoding a No. US2002009455A1e1 PROST 03  
FILE REFERENCE: 51831AUSM1

CURRENT APPLICATION NUMBER: US/09/838,785

PRIOR FILING DATE: 2001-04-20

PRIOR APPLICATION NUMBER: 60/200,065

NUMBER OF SEQ ID NOS: 26

SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1  
LENGTH: 3320  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (282) .. (1943)  
US-09-838-785-1

Alignment Scores:  
Pred. No.: 27.5 Length: 3320  
Score: 50.00 Matches: 9  
Percent Similarity: 65.00% Conservative: 4  
Best Local Similarity: 45.00% Mismatches: 7  
Query Match: 34.48% Indels: 0  
DB: 10 Gaps: 0

US-09-987-190-2 (1-30) x US-09-838-785-1 (1-3320)

Qy 2 TySerleupProgluLeuApyTyrgluPheserAlaThrgluProTyrlleserGlyln 21  
Db 807 TACTCTGCTGCTGCATTGACGTGGACACCACTGCTGCCCTTACTGGGACCCAG 866

Search completed: April 8, 2003, 05:00:28  
Job time : 72.1651 secs

GenCore version 5.1.4 p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 9, 2003, 13:51:49 ; Search time 14.8624 Seconds  
(without alignments)  
415.910 Million cell updates/sec

Title: US-09-987-190-2  
Perfect score: 145  
Sequence: 1 KXSLPELDYFSGATEPYISGQINEXYTX 30

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

rchd: 671580 seqs, 206047115 residues  
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_21.\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriophage:\*  
17: sp\_archaeal:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	107	73.8	237	3	P79022
2	101	69.7	207	2	Q9F9R1
3	101	69.7	210	2	Q9AM00
4	91	62.8	189	2	Q59673
5	80	55.2	32	8	Q9T2Q1
6	80	55.2	201	5	Q8TA63
7	77	53.1	200	2	Q9APY3
8	76	52.4	216	10	Q9FY32
9	75	51.7	216	3	Q8XIN7
10	74	51.0	140	4	Q96AM7
11	74	51.0	165	4	Q9UG59
12	74	51.0	218	11	Q8VEM5
13	74	51.0	232	4	Q96RE6
14	74	51.0	231	10	Q65324
15	74	51.0	231	10	Q947R3
16	73	50.3	203	2	Q9MWG7

17	73	50.3	206	3	Q96UT6	Q96UT6 candida alb
18	72	49.7	211	17	Q96Y84	Q96Y84 sulfolobus
19	71	49.0	203	2	Q93K55	Q93K55 frankia sp.
20	71	49.0	205	10	Q9STB5	Q9STB5 hevea bras
21	71	49.0	205	10	Q9FSJ2	Q9FSJ2 hevea bras
22	71	49.0	208	3	Q96UR1	Q96UR1 colletoic
23	71	49.0	213	2	Q9X6N3	Q9X6N3 streptomye
24	71	49.0	220	3	Q74200	Q74200 pneumocyst
25	71	49.0	224	10	Q9FY33	Q9FY33 digitalis 1
26	71	49.0	224	13	Q9DDJ1	Q9DDJ1 gallus gall
27	71	49.0	224	13	Q9Y334	Q9Y334 gallus gall
28	71	49.0	227	10	Q9M532	Q9M532 euphorbia e
29	69	47.6	223	3	Q9E945	Q9E945 emericella.
30	68	46.9	201	2	Q9F326	Q9F326 staphylococ
31	68	46.9	203	16	Q9PAA4	Q9PAA4 xylella fas
32	68	46.9	206	3	Q96WZ8	Q96WZ8 phanerocha
33	68	46.9	215	16	Q9X469	Q9X469 streptomye
34	68	46.9	231	10	P93606	P93606 triticum ae
35	68	46.9	240	10	Q82584	Q82584 zantedesch
36	67	46.2	202	2	Q9KW85	Q9KW85 vibrio vuln
37	67	46.2	211	16	Q9RUV2	Q9RUV2 deinococcus
38	67	46.2	218	5	Q9NB66	Q9NB66 callinectes
39	67	46.2	228	10	Q9SM64	Q9SM64 prunus pers
40	67	46.2	235	10	Q43273	Q43273 zea mays (m
41	66	45.5	25	8	Q9T2Q0	Q9T2Q0 solanum tub
42	66	45.5	206	3	Q9Y773	Q9Y773 paxillus in
43	66	45.5	207	16	Q829X6	Q829X6 yersinia pe
44	65	44.8	224	10	Q82571	Q82571 triticum ae
45	65	44.8	231	10	Q96185	Q96185 triticum ae

## ALIGNMENTS

## RESULT 1

P79022	PRELIMINARY;	PRT;	227 AA.
AC P79022;			
DT 01-MAY-1997 (TREMBLrel. 03, Created)			
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)			
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE Superoxide dismutase (EC 1.15.1.1) precursor.			
GN MNSOD.			
OS Candida sp. HN95.			
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;			
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.			
OX NCBI_TaxID=159257;			
RN [1]			
RP SEQUENCE FROM N.A.			
RA Choi S.-Y., Nam Y.-S., Hong Y.M.;			
RT "Molecular cloning and characterization of Mn-superoxide dismutase			
RL gene from Candida sp.";			
CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE			
CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).			
CC -1- CATALYTIC ACTIVITY: 2 SUPEROXIDE + 2 H(+) = O(2) + H(2)O(2).			
CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE			
CC FAMILY			
DR EMBL; Y11596; CAA72335.1; --			
DR HSSP; P04179; IAP6.			
DR InterPro; IPR001189; SODismutase.			
DR Pfam; PF00081; sode; 1.			
DR Pfam; PF02777; sode; C; 1.			
DR ProDom; PD000475; SODismutase; 1.			
DR PROSITE; PS00088; SOD_MN; 1.			
KW Oxidoreductase; Signal.			
FT STGNL			
FT CHAIN			
SQ SEQUENCE 227 AA; 25227 MW; 29E7FEBC7743DB0 CRC64;			
Query Match	73.8%;	Score 107;	DB 3; Length 227;
Best Local Similarity	71.4%;	Pred. No. 2.1e-08;	
Matches	20;	Conservative	4; Mismatches 4; Indels 0; Gaps 0;





RESULT 5

Q9T201 PRELIMINARY; PRT; 32 AA.  
AC Q9T201;  
DT 01-MAY-2000 (TREMblrel. 13, Created)  
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)  
DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
DE Manganease superoxide dismutase, MN-SOD (EC 1.15.1.1) (Fragment).  
OS Pinnus sylvestris (Scots pine).  
OG Mitochondrion.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.  
OX NCBI\_Taxid=3349;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=95072927; PubMed=7981961;  
RA Streller S., Kromer S., Winkler G.;  
"Isolation and purification of mitochondrial Mn-superoxide dismutase from the gymnosperm Pinnus sylvestris L.";  
Plant Cell Physiol. 35:859-867(1994).  
CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: 2 SUPEROXIDE + 2 H(+) = O(2) + H(2)O(2).  
CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE FAMILY.  
DR InterPro: IPR001189; SODismutase.  
DR Pfam: PF00081; sodfe; 1.  
DR Prodom: PD000475; SODismutase; 1.  
KM Manganease; Oxidoreductase.  
SQ SEQUENCE 32 AA; 3681 MW; 7A6826DC1427092C CRC64;

Query Match 55.2%; Score 80; DB 8; Length 32;  
Best Local Similarity 57.7%; Pred. No. 3.4e-05;  
Matches 15; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 2 YSLPELDYFSATEPYISQINEIXYT 27  
:|||||:|||||:  
DB 4 FSLPELDYFSATEPYISQINEIXYT 29

RESULT 6

Q8TA63 PRELIMINARY; PRT; 201 AA.  
AC Q8TA63;  
DT 01-JUN-2002 (TREMblrel. 21, Created)  
DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)  
DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
DE Superoxide dismutase like protein.  
OS SOD-1.  
OS Pinnus japonicus (Kuruma prawn).  
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
OC Eumalacostraca; Eucarida; Decapoda; Dendrobranchiata; Penaeoidea;  
OC Penaeidae; Marsupenaeus.  
OX NCBI\_Taxid=27405;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Soma G., Inagawa H., Nishizawa T., Honda T., Nomura Y., Uenobe M.,  
Takashi Y., Itami T., Yokomizo Y.;  
"Kuruma shrimp cDNA similar to superoxide dismutase protein.";  
Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB079877; BAB85211.1; -  
SQ SEQUENCE 201 AA; 21485 MW; 3317CBEF5FDA9A52 CRC64;

Query Match 55.2%; Score 80; DB 5; Length 201;  
Best Local Similarity 55.6%; Pred. No. 0.00027;  
Matches 15; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 2 YSLPELDYFSATEPYISQINEIXYT 28  
:|||||:|||||:  
DB 87 HALPLKTYDPALEPHISGIMELIHT 113

RESULT 7

Q9APY3 PRELIMINARY; PRT; 200 AA.  
AC Q9APY3;  
DT 01-JUN-2001 (TREMblrel. 17, Created)  
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)  
DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
DE Manganease superoxide dismutase (EC 1.15.1.1).  
GN SODA.  
OS Corynebacterium melassecola, and  
OS Corynebacterium glutamicum (Brevibacterium flavum).  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Corynebacterineae; Corynebacteriaceae;  
OC Corynebacterium.  
OX NCBI\_Taxid=41643, 1718;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC SPECIES=C.melassecola; STRAIN=ATCC17965;  
RX MEDLINE=21101811; PubMed=11157941;  
RA Merkamm M., Guyonvarch A.;  
"Cloning of the soda gene from Corynebacterium melassecola and Role of Superoxide Dismutase in Cellular Viability.";  
J. Bacteriol. 183:1284-1295(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC SPECIES=C.glutamicum; STRAIN=ATCC 13032;  
RA Nakagawa S., Mizoguchi H., Ando S., Hayaehi M., Hattori M., Shiba T.,  
Sakaki Y., Yokoi H., Ozaki A.;  
"SOD of Corynebacterium glutamicum ATCC 13032.";  
Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: 2 SUPEROXIDE + 2 H(+) = O(2) + H(2)O(2).  
CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE FAMILY.  
DR EMBL; AF236111; AKK01490.1; -  
DR EMBL; AB055218; BAB62412.1; -  
DR HSSP; P17670; 1IDS.  
DR InterPro: IPR001189; SODismutase.  
DR Pfam: PF00081; sodfe; 1.  
DR Pfam: PF02277; sodfe; C; 1.  
DR Prodom: PD000475; SODismutase; 1.  
DR PROSITE; PS00088; SOD\_MN; 1.  
KM Oxidoreductase.  
SQ SEQUENCE 200 AA; 22101 MW; D236A2C3F291CE13 CRC64;

Query Match 53.1%; Score 77; DB 2; Length 200;  
Best Local Similarity 48.1%; Pred. No. 0.00077;  
Matches 13; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 2 YSLPELDYFSATEPYISQINEIXYT 28  
:|||||:|||||:  
DB 4 YELPELDYADALEPHIAEIMELHNS 30

RESULT 8

Q9FY32 PRELIMINARY; PRT; 216 AA.  
AC Q9FY32;  
DT 01-MAR-2001 (TREMblrel. 16, Created)  
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)  
DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
DE Manganease superoxide dismutase (EC 1.15.1.1) (Fragment).  
GN MN-SOD.  
OS Digitalis lanata (Foxglove).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliopsida; Eudicotyledons; core eudicots;  
OC Asteridae; euasterids I; Lamiales; Veronicaceae; Digitalis.  
OX NCBI\_Taxid=49450;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Gruner B., Reva V.A., Mueller-Urli F.;  
"Stress-induction of Mn-dependent SOD and CAT in Digitalis lanata

RT (shrh.) during somatic embryogenesis." ;  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE  
 CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: 2 SUPEROXIDE + 2 H(+) = O(2) + H(2)O(2).  
 CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE  
 CC FAMILY.  
 DR EMBL; AJ278664; CAC05260.1; -.  
 DR HSSP; P04179; IAP6.  
 DR InterPro; IPR001189; SODismutase.  
 DR Pfam; PF00081; sodfe; 1.  
 DR Pfam; PF02777; sodfe\_C; 1.  
 DR Prodom; PD000475; SODismutase; 1.  
 DR PROSITE; PS00088; SOD\_MN; 1.  
 KM Oxidoreductase.  
 FT NON TER 1 1  
 SQ SEQUENCE 216 AA; 23923 MW; 5CDD66EF9FC73077 CRC64;  
 Query Match 52.4%; Score 76; DB 10; Length 216;  
 Best Local Similarity 51.9%; Pred. No. 0.0012;  
 Matches 14; Conservative 7; Mismatches 6; Indels 0; Gaps 0;  
 Db 1 KXSLPDLPEFSATEPYISQGINEIXYT 27  
 15 KXSLPDLPEYDYGALPEHISQIMQIHH 41  
 RESULT 9  
 O8XIN7 PRELIMINARY; PRT; 216 AA.  
 AC O8XIN7;  
 DT 01-MAR-2002 (TREMBlrel. 20, Created)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Superoxide dismutase.  
 GN SOD1.  
 OS Blumeria graminis.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomycetes;  
 CC Erysiphales; Erysiphaceae; Blumeria.  
 OX NCBI\_TaxId=34373;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RA Zhang Z., Gurr S.J.;  
 RT "Isolation and characterization of a Blumeria graminis superoxide  
 RT dismutase gene involved in antioxidant action in barley/powdery mildew  
 RT interactions." ;  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF227338; AAJ56985.1; -.  
 DR InterPro; IPR001189; SODismutase.  
 DR Pfam; PF00081; sodfe; 1.  
 DR Pfam; PF02777; sodfe\_C; 1.  
 DR Prodom; PD000475; SODismutase; 1.  
 DR PROSITE; PS00088; SOD\_MN; UNKNOWN 1.  
 DR SEQUENCE 216 AA; 24194 MW; 20EB6F1A027AA9C CRC64;  
 Query Match 51.7%; Score 75; DB 3; Length 216;  
 Best Local Similarity 50.0%; Pred. No. 0.0017;  
 Matches 14; Conservative 5; Mismatches 9; Indels 0; Gaps 0;  
 Db 1 KXSLPDLPEFSATEPYISQGINEIXYT 28  
 4 KYNIPPLPYAYDALEPHISQIMINHHT 31  
 RESULT 10  
 O96AM7 PRELIMINARY; PRT; 140 AA.  
 AC O96AM7;  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Hypothetical 15.7 kDa protein.  
 OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=BRAIN;  
 RA Strauberg R.;  
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE  
 CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: 2 SUPEROXIDE + 2 H(+) = O(2) + H(2)O(2).  
 CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE  
 CC FAMILY.  
 DR EMBL; BC016934; AAH16934.1; -.  
 DR InterPro; IPR001189; SODismutase.  
 DR Pfam; PF00081; sodfe; 1.  
 DR Prodom; PD000475; SODismutase; 1.  
 KM Hypothetical protein; Oxidoreductase.  
 FT NON TER 1 1  
 SQ SEQUENCE 140 AA; 15745 MW; 499AC0467EF5C451 CRC64;  
 Query Match 51.0%; Score 74; DB 4; Length 140;  
 Best Local Similarity 42.9%; Pred. No. 0.0015;  
 Matches 12; Conservative 10; Mismatches 6; Indels 0; Gaps 0;  
 Db 1 KXSLPDLPEFSATEPYISQGINEIXYT 28  
 25 KXSLPDLPEYDYGALPEHISQIMQIHH 52  
 RESULT 11  
 O9UG59 PRELIMINARY; PRT; 165 AA.  
 AC O9UG59;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Hypothetical 18.6 kDa protein (fragment).  
 GN DKFZP64W2422.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxId=9606;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RA Mambour R., Heubner D., Mewes H.W., Gassenhuber J., Wiemann S.;  
 RC TISSUE=BRAIN;  
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE  
 CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: 2 SUPEROXIDE + 2 H(+) = O(2) + H(2)O(2).  
 CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE  
 CC FAMILY.  
 DR EMBL; AL050388; CAB62521.1; -.  
 DR HSSP; P04179; IAP6.  
 DR InterPro; IPR001189; SODismutase.  
 DR Pfam; PF00081; sodfe; 1.  
 DR Prodom; PD000475; SODismutase; 1.  
 DR Hypothetical protein; Oxidoreductase.  
 FT NON TER 1 1  
 SQ SEQUENCE 165 AA; 18605 MW; C7F2C078AD127205 CRC64;  
 Query Match 51.0%; Score 74; DB 4; Length 165;  
 Best Local Similarity 42.9%; Pred. No. 0.0018;  
 Matches 12; Conservative 10; Mismatches 6; Indels 0; Gaps 0;  
 Db 1 KXSLPDLPEFSATEPYISQGINEIXYT 28  
 50 KXSLPDLPEYDYGALPEHISQIMQIHH 77  
 RESULT 12  
 O8VEM5 PRELIMINARY; PRT; 218 AA.  
 ID O8VEM5



DR PROSITE; PS00088; SOD\_MN; UNKNOWN\_1.  
KM Oxidoreductase.  
SQ SEQUENCE 231 AA; 25454 MW; 9805F5D3B949BEC CRC64;

Query Match 51.0%; Score 74; DB 10; Length 231;  
Best Local Similarity 50.0%; Pred. No. 0.0026;  
Matches 13; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 2 YSLPELDYEFSAEPYISQINEIXY 27  
DB 31 FTLPDLPIYDALPRAISGEIMQH 56

## RESULT 16

Q9WMG7 PRELIMINARY; PRT; 203 AA.

AC Q9WMG7; (TREMBlrel. 12, Created)  
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE Manganese superoxide dismutase.

Pseudomonas syringae (pv. syringae);  
Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
Pseudomonas.

NCBI\_TaxID=321;

SEQUENCE FROM N.A.

STRAIN=B728A.

Kim Y.C., Miller C.D., Anderson A.J.,

"Transcriptional regulation and mutational analysis of genes encoding

iron- and manganese-superoxide dismutases from the phytopathogenic

bacterium, Pseudomonas syringae pv. syringae B728A."

Submitted (JAN-1999) to the EMBL/Genbank/DBJ databases.

- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE

CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).

- CATALYTIC ACTIVITY: 2 SUPEROXIDE + 2 H(+) = O(2) + H(2)O(2).

- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE

FAMILY.

EMBL; AF121078; AAD24796.1; -.

HSSP; P00448; IYEW.

InterPro; IPR001189; SODismutase.

Pfam; PF00081; sodfe, 1.

Pfam; PF02777; sodfe, C, 1.

ProDom; PD000475; SODismutase; 1.

PROSITE; PS00088; SOD\_MN; 1.

Oxidoreductase.

SEQUENCE 203 AA; 22628 MW; BAF6DA34642F155B CRC64;

Query Match 50.3%; Score 73; DB 2; Length 203;

Best Local Similarity 51.9%; Pred. No. 0.0032;

Matches 14; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 2 YSLPELDYEFSAEPYISQINEIXY 28  
DB 3 YTLPLPYAYDALPRAISGEIMQH 29

## RESULT 17

Q96UT6 PRELIMINARY; PRT; 206 AA.

AC Q96UT6; (TREMBlrel. 19, Created)

DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE Manganese-containing superoxide dismutase.

SOD3.

Candida albicans (Yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; mitosporic Saccharomycetales; Candida.

NCBI\_TaxID=5476;

SEQUENCE FROM N.A.

RA Lamotte C., LeMay J.-D., Deslauniers N., Bourbonnais Y.;  
RT "Candida albicans expresses an unusual cytoplasmic manganese-  
containing superoxide dismutase SOD3 upon entry and during stationary

phase."

RT phase."

J. Biol. Chem. 0:0-0(2001).

- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE

CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).

- CATALYTIC ACTIVITY: 2 SUPEROXIDE + 2 H(+) = O(2) + H(2)O(2).

- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE

FAMILY.

EMBL; AF16340; AAL09560.1; -.

InterPro; IPR001189; SODismutase.

Pfam; PF00081; sodfe, 1.

Pfam; PF02777; sodfe, C, 1.

ProDom; PD000475; SODismutase; 1.

PROSITE; PS00088; SOD\_MN; UNKNOWN\_1.

Oxidoreductase.

SEQUENCE 206 AA; 22734 MW; 303FA3503417F332 CRC64;

Query Match 50.3%; Score 73; DB 3; Length 206;

Best Local Similarity 52.0%; Pred. No. 0.0033;

Matches 13; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 1 KYSLPELDYEFSAEPYISQINEI 25  
DB 7 KISLPKIDWALDALEPYISKEINDL 31

## RESULT 18

Q96Y84 PRELIMINARY; PRT; 211 AA.

AC Q96Y84; (TREMBlrel. 19, Created)

DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE Putative superoxide dismutase.

ST2283.

OS Sulfolobus tokodaii.

OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;

Sulfolobus.

NCBI\_TaxID=111955;

SEQUENCE FROM N.A.

STRAIN=JCM 10545 / 7;

PdbMed=11572479;

Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,

Ra Sekine M., Baba S.-I., Anka A., Kosugi H., Hosoyama A., Fukui S.,

Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,

Yoshizawa T., Tanaka T., Kudo Y., Yamazaki J., Kushiida N., Oguchi A.,

Aoki K.-I., Maeda S., Yanagii M., Nishimura M., Yamagishi A.,

Oshima T., Kikuchi H.;

"Complete genome sequence of an aerobic thermoacidophilic

Crenarchaeon, Sulfolobus tokodaii strain7."

DNA Res. 8:123-140(2001).

EMBL; AP000989; BAB67393.1; -.

InterPro; IPR001189; SODismutase.

Pfam; PF00081; sodfe, 1.

Pfam; PF02777; sodfe, C, 1.

ProDom; PD000475; SODismutase; 1.

Hypothetical protein; Complete proteome.

SEQUENCE 211 AA; 24302 MW; 0EEAC4A76938002 CRC64;

Query Match 49.7%; Score 72; DB 17; Length 211;

Best Local Similarity 51.9%; Pred. No. 0.0048;

Matches 14; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 KYSLPELDYEFSAEPYISQINEIXY 27

DB 9 KYELPLPYVADLPIYSKIDIVHY 35

## RESULT 19

Q93K55

ID Q93K55 PRELIMINARY; PRT; 203 AA.  
 AC Q93K55;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Iron superoxide dismutase (EC 1.15.1.1).  
 GN SOD.  
 OS Frankia sp. ACN14a-ter.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Frankineae; Frankiaceae; Frankia.  
 OX NCBI\_TaxID=123227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ACN14a-ter;  
 RX MEDLINE=21360358; PubMed=11467730;  
 RA Hammad Y., Marechal J., Courmoyer B., Normand P., Dome A.M.;  
 RT "Modification of the protein expression pattern induced in the nitrogen-fixing actinomycete Frankia sp. strain ACN14a-ter by root exudates of its symbiotic host Alnus glutinosa and cloning of the sodF gene.";  
 RL Can. J. Microbiol. 47:541-547(2001).  
 CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: 2 SUPEROXIDE + 2 H(+) = O(2) + H(2)O(2).  
 CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE FAMILY.  
 CC EMBL: AJ277540; CAC60251.1; -.  
 CC InterPro: IPR001189; SODismutase.  
 DR Pfam: PF02777; sodfe\_C; 1.  
 DR Pfam: PF02777; sodfe\_C; 1.  
 DR ProDom: PD000475; SODismutase; 1.  
 KW Oxidoreductase.  
 SQ SEQUENCE 203 AA; 22141 MW; AA71ECF9D524377 CRC64;  
 QY 2 YSLPELDYFSATEPYISGOINEIXY 27  
 Db 4 YSLPDLKDYGALEPALSIGQIVLHH 29

RESULT 20  
 Q9STB5 PRELIMINARY; PRT; 205 AA.  
 Q9STB5;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE MnSOD (Fragment).  
 GN SOD.  
 OS Hevea brasiliensis (Para rubber tree).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC euroside 1; Malpighiales; Euphorbiaceae; Hevea.  
 OX NCBI\_TaxID=3981;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Wagner S., Breiteneder H.;  
 RT "Cloning and molecular characterization of MnSOD, a cross-reactive allergen from Hevea brasiliensis latex."  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: 2 SUPEROXIDE + 2 H(+) = O(2) + H(2)O(2).  
 CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE FAMILY.  
 CC EMBL: AJ249148; CAB53458.1; -.  
 CC HSSP: P04179; IAP6.  
 DR InterPro: IPR001189; SODismutase.  
 DR Pfam: PF00081; sodfe\_1.  
 DR Pfam: PF02777; sodfe\_C; 1.

DR ProDom: PD000475; SODismutase; 1.  
 DR PROSITE: PS00088; SOD\_MN; 1.  
 KW Oxidoreductase.  
 FT NON TER 1 1  
 SQ SEQUENCE 205 AA; 22915 MW; 7E6889364E63BA21 CRC64;  
 QY 2 YSLPELDYFSATEPYISGOINEIXY 27  
 Db 3 FSLPDLPYDGALEPALSIGIMQVLH 28

Query Match 49.0%; Score 71; DB 10; Length 205;  
 Best Local Similarity 46.2%; Pred. No. 0.0067;  
 Matches 12; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

RESULT 21  
 Q9FSJ2 PRELIMINARY; PRT; 205 AA.  
 Q9FSJ2;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE IGE-binding protein MnSOD (fragment).  
 GN SOD.  
 OS Hevea brasiliensis (Para rubber tree).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC euroside 1; Malpighiales; Euphorbiaceae; Hevea.  
 OX NCBI\_TaxID=3981;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=RRIM 600; TISSUE=LEAF;  
 RX MEDLINE=21068908; PubMed=11167363;  
 RA Riba H.P., Chen Z., Roznyek P., Cremer R.;  
 RT "Allergenicity of rhey b 10 (Manganese superoxide dismutase).";  
 RL Allergy 56:85-86(2001).  
 CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: 2 SUPEROXIDE + 2 H(+) = O(2) + H(2)O(2).  
 CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE FAMILY.  
 CC EMBL: AJ289158; CAC13961.1; -.  
 CC HSSP: P04179; IAP6.  
 CC InterPro: IPR001189; SODismutase.  
 DR Pfam: PF00081; sodfe\_1.  
 DR Pfam: PF02777; sodfe\_C; 1.  
 DR ProDom: PD000475; SODismutase; 1.  
 DR PROSITE: PS00088; SOD\_MN; 1.  
 KW IGE-binding protein; Oxidoreductase.  
 FT NON TER 1 1  
 SQ SEQUENCE 205 AA; 22915 MW; DC0F9D3EB3E97A3B CRC64;  
 QY 2 YSLPELDYFSATEPYISGOINEIXY 27  
 Db 3 FSLPDLPYDGALEPALSIGIMQVLH 28

Query Match 49.0%; Score 71; DB 10; Length 205;  
 Best Local Similarity 46.2%; Pred. No. 0.0067;  
 Matches 12; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

RESULT 22  
 Q96UR1 PRELIMINARY; PRT; 208 AA.  
 Q96UR1;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Manganese superoxide dismutase.  
 GN SOD2.  
 OS Colletotrichum graminicola (Anthracnose fungus) (Glomerella graminicola).  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;

OC Sordariomycetes incertae sedis; Phyllachorales; Phyllachoraceae;  
 OC Glomerella.  
 OX NCBI\_TaxID=31870;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Fang G., Hanau R., Vallancourt L.,  
 RT "Molecular characterization of the SOD2 gene, encoding a manganese  
 type superoxide dismutase in the plant-pathogenic fungus  
 RT Colletotrichum graminicola."  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE  
 CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: 2 SUPEROXIDE + 2 H(+) = O(2) + H(2)O(2).  
 CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE  
 CC FAMILY.  
 DR EMBL: AF430836; AL27457.1; -.  
 DR InterPro: IPR001189; SODismutase.  
 DR Pfam: PF00081; sode; 1.  
 DR Pfam: PF02777; sode\_C; 1.  
 DR ProDom: PD000475; SODismutase; 1.  
 DR PROSITE: PS00088; SOD\_MN; UNKNOWN\_1.  
 DR Oxidoreductase.  
 SEQUENCE 208 AA; 23026 MW; COB22E4AE6E225E CRC64;

Query Match 49.0%; Score 71; DB 3; Length 208;  
 Best Local Similarity 46.4%; Pred. No. 0.0068;  
 Matches 13; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 KYSLPELDYFSATEPYISQINEXYT 28  
 Db 5 QYILPELPYADALEPHISAQIMELHHS 32

## RESULT 23

09X6N3 PRELIMINARY; PRT; 213 AA.  
 ID Q9X6N3;  
 AC Q9X6N3;  
 DT 01-NOV-1999 (TREMblrel. 12, Created)  
 DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE Re-Zn-superoxide dismutase.  
 GN SODF.  
 OS Streptomyces griseus.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1911;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20435297; PubMed=10978523;  
 RA Kim J.-S., Jang J.-H., Lee J.-W., Kang S.-O., Kim K.-S., Lee J.K.,  
 RT "Identification of cis site involved in nickel-responsive  
 transcriptional repression of sode gene coding for Fe- and Zn-  
 containing superoxide dismutase of Streptomyces griseus."  
 RL Blochm. Biophys. Acta 1493:200-207(2000).  
 CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE  
 CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: 2 SUPEROXIDE + 2 H(+) = O(2) + H(2)O(2).  
 CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE  
 CC FAMILY.  
 DR EMBL: AF141866; AAD30139.1; -.  
 DR HSP: P17670; 1IDS.  
 DR InterPro: IPR001189; SODismutase.  
 DR Pfam: PF00081; sode; 1.  
 DR Pfam: PF02777; sode\_C; 1.  
 DR ProDom: PD000475; SODismutase; 1.  
 DR PROSITE: PS00088; SOD\_MN; 1.  
 DR Oxidoreductase.  
 SEQUENCE 213 AA; 23476 MW; 57989919447C7583 CRC64;

Query Match 49.0%; Score 71; DB 2; Length 213;  
 Best Local Similarity 50.0%; Pred. No. 0.007;  
 Matches 13; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 2 YSLPELDYFSATEPYISQINEXYT 27  
 Db 4 YTLPELPYADALEPHISAQIMELHHS 29

## RESULT 24

074200 PRELIMINARY; PRT; 220 AA.  
 ID 074200;  
 AC 074200;  
 DT 01-NOV-1998 (TREMblrel. 08, Created)  
 DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE Manganese superoxide dismutase precursor.  
 GN MNSOD.  
 OS Pneumocystis carinii.  
 OC Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes; Pneumocystidaceae;  
 OC Pneumocystis.  
 OX NCBI\_TaxID=4754;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Denis C.M., Guyot K., Del-Cas E., Camus D., Odberg-Ferragut C.;  
 RT "Mn SOD gene from rat derived P. carinii."  
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE  
 CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: 2 SUPEROXIDE + 2 H(+) = O(2) + H(2)O(2).  
 CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE  
 CC FAMILY.  
 DR EMBL: AF036321; AAC24764.1; -.  
 DR HSP: P04179; 1ABW.  
 DR InterPro: IPR001189; SODismutase.  
 DR InterPro: IPR000834; Zn\_cardoepc.  
 DR Pfam: PF00081; sode; 1.  
 DR Pfam: PF02777; sode\_C; 1.  
 DR ProDom: PD000475; SODismutase; 1.  
 DR PROSITE: PS00133; CARBOXYPEPT\_ZN\_2; UNKNOWN\_1.  
 DR PROSITE: PS00088; SOD\_MN; 1.  
 DR Oxidoreductase; Signal.  
 FT SIGNAL 1 27  
 FT CHAIN 28 220  
 FT POTENTIAL.  
 SEQUENCE 220 AA; 25869 MW; 73B1F1C98929E18 CRC64;

Query Match 49.0%; Score 71; DB 3; Length 220;  
 Best Local Similarity 44.4%; Pred. No. 0.0072;  
 Matches 12; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 1 KYSLPELDYFSATEPYISQINEXYT 27  
 Db 24 KHVLPSPDYQALEPYSADLIELHY 50

## RESULT 25

09FY33 PRELIMINARY; PRT; 224 AA.  
 ID 09FY33;  
 AC 09FY33;  
 DT 01-MAR-2001 (TREMblrel. 16, Created)  
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE Manganese superoxide dismutase (EC 1.15.1.1) (Fragment).  
 GN MN-SOD.  
 OS Digitalis lanata (Foxglove).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids I; Lamiales; Veroniceae; Digitalis.  
 OX NCBI\_TaxID=49450;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Gruner B., Reva V.A., Mueller-Urli F.;  
 RT "Stress-induction of Mn-dependent SOD and CAT in Digitalis lanata  
 (Bhrh.) during somatic embryogenesis."  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE  
 CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).

CC -1- CATALYTIC ACTIVITY: 2 SUPEROXIDE + 2 H(+) = O(2) + H(2)O(2).  
 CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE  
 CC FAMILY.  
 DR EMBL; AJ276863; CAC05259.1; -.  
 DR HSSP; P04179; IABM.  
 DR InterPro; IPR001189; SODismutase.  
 DR Pfam; PF00081; sodfe\_1.  
 DR Pfam; PF02777; sodfe\_C; 1.  
 DR ProDom; PD000475; SODismutase; 1.  
 DR PROSITE; PS00088; SOD\_MN; 1.  
 KW Oxidoreductase.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 224 AA; 24757 MW; DD569F3F62C66E08 CRC64;  
 Query Match 49.0%; Score 71; DB 10; Length 224;  
 Best Local Similarity 46.2%; Pred. No. 0.0074;  
 Matches 12; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

RESULT 26  
 Q9DDJ1 PRELIMINARY; PRT; 224 AA.

ID Q9DDJ1  
 AC Q9DDJ1; PRELIMINARY; PRT; 224 AA.  
 DT 01-MAR-2001 (T-EMBLrel. 16, Created)  
 DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)  
 DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)  
 DE Manganese-containing superoxide dismutase precursor (EC 1.15.1.1).  
 GN MNSOD.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=HEART;  
 RA Bu Y.Q., Luo X.G., Li S.F., Lu C., Li Y.W., Kuang X., Liu B., Li J.F.,  
 RA Yu S.X.;  
 RT "Cloning and Sequence Analysis of Manganese-containing Superoxide  
 RT Dismutase (Mnsod) cDNA of Chicken."  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE  
 CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: 2 SUPEROXIDE + 2 H(+) = O(2) + H(2)O(2).  
 CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE  
 CC FAMILY.  
 CC EMBL; AF329270; AAG46055.1; -.  
 DR HSSP; P04179; IABM.  
 DR InterPro; IPR001189; SODismutase.  
 DR Pfam; PF00081; sodfe\_1.  
 DR Pfam; PF02777; sodfe\_C; 1.  
 DR ProDom; PD000475; SODismutase; 1.  
 DR PROSITE; PS00088; SOD\_MN; 1.  
 KW Oxidoreductase; Signal.  
 FT SIGNAL 1 26  
 FT CHAIN 27 224  
 SQ SEQUENCE 224 AA; 24946 MW; 119211AC7266687F CRC64;

Query Match 49.0%; Score 71; DB 13; Length 224;  
 Best Local Similarity 39.3%; Pred. No. 0.0074;  
 Matches 11; Conservative 11; Mismatches 6; Indels 0; Gaps 0;

QY 1 KYSLPELDYEFSAPEYISQINEIXYT 28  
 Db 27 KHTLPDLPDYGALPEPHISAIMQLHHS 54

RESULT 27

Q90Y34  
 ID Q90Y34 PRELIMINARY; PRT; 224 AA.  
 AC Q90Y34;  
 DT 01-DEC-2001 (T-EMBLrel. 19, Created)  
 DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)  
 DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)  
 DE Mnsod.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kim H., You S., Foster L.K., Parrie J., Foster D.N.;  
 RT "Expression of antioxidant genes in primary and immortal chicken  
 RT embryo fibroblast cells."  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE  
 CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: 2 SUPEROXIDE + 2 H(+) = O(2) + H(2)O(2).  
 CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE  
 CC FAMILY.  
 CC EMBL; AF299388; AAK97214.1; -.  
 DR InterPro; IPR001189; SODismutase.  
 DR Pfam; PF00081; sodfe\_1.  
 DR Pfam; PF02777; sodfe\_C; 1.  
 DR ProDom; PD000475; SODismutase; 1.  
 DR PROSITE; PS00088; SOD\_MN; UNKNOWN\_1.  
 KW Oxidoreductase.  
 SQ SEQUENCE 224 AA; 24770 MW; A1B570FC736BD3DD CRC64;  
 Query Match 49.0%; Score 71; DB 13; Length 224;  
 Best Local Similarity 39.3%; Pred. No. 0.0074;  
 Matches 11; Conservative 11; Mismatches 6; Indels 0; Gaps 0;

QY 1 KYSLPELDYEFSAPEYISQINEIXYT 28  
 Db 27 KHTLPDLPDYGALPEPHISAIMQLHHS 54

RESULT 28  
 Q9M532 PRELIMINARY; PRT; 237 AA.  
 ID Q9M532  
 AC Q9M532;  
 DT 01-OCT-2000 (T-EMBLrel. 15, Created)  
 DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)  
 DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)  
 DE Manganese superoxide dismutase.  
 OS Euphorbia esula (leafy spurge).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC Eurosid 1; Malpighiales; Euphorbiaceae; Euphorbia.  
 OX NCBI\_TaxID=3993;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=UNDERGROUND ADVENTITIOUS BUDS;  
 RA Anderson J.V., Horvath D.P.;  
 RT "Identification of mRNAs expressed in underground adventitious buds of  
 RT Euphorbia esula (leafy spurge)."  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE  
 CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: 2 SUPEROXIDE + 2 H(+) = O(2) + H(2)O(2).  
 CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE  
 CC FAMILY.  
 CC EMBL; AF242310; AAF65768.1; -.  
 DR HSSP; P04179; IAP6.  
 DR InterPro; IPR001189; SODismutase.  
 DR Pfam; PF00081; sodfe\_1.  
 DR Pfam; PF02777; sodfe\_C; 1.  
 DR ProDom; PD000475; SODismutase; 1.  
 DR PROSITE; PS00088; SOD\_MN; 1.







Best Local Similarity 50.0%; Pred. No. 0.02;  
Matches 13; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 2 YSLPELDYFSATEPYISGOINEIXY 27  
Db 4 YTLPELDYDYSALPAVISPETIELHH 29

## RESULT 34

P93606 PRELIMINARY; PRT; 231 AA.

AC P93606; 01-MAY-1997 (TREMBlrel. 03, Created)  
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE Manganese superoxide dismutase (EC 1.15.1.1).  
GN SOD3.2.  
OS Triticum aestivum (Wheat).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
OC Triticaceae; Triticum.  
NCBI\_TaxId=4565;  
[1]

SEQUENCE FROM N.A.  
Wu G., Robertson A., Wilen R., Gust L.;  
"Isolation and characterization of two cDNAs (Accession Nos. U72212 and U73172) encoding mitochondrial manganese superoxide dismutases in wheat (PGR97-023)."  
RL Plant Physiol. 113:664-664(1997).  
CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: 2 SUPEROXIDE + 2 H(+) = O(2) + H(2)O(2).  
CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE FAMILY.

DR EMBL: U73172; AAB68036.1; -.  
DR HSSP: P04179; IAB6.  
DR InterPro: IPR001189; SODismutase.  
DR Pfam: PF00081; sode; 1.  
DR Pfam: PF02777; sode; C; 1.  
DR ProDom: PD000475; SODismutase; 1.  
DR PROSITE: PS00088; SOD\_MN; 1.  
KM Oxidoreductase.

SQ SEQUENCE 231 AA; 25299 MW; 01CSB547E2F8FFD0 CRC64;

Query Match 46.9%; Score 68; DB 10; Length 231;  
Best Local Similarity 42.3%; Pred. No. 0.022; Mismatches 7; Indels 0; Gaps 0;

QY 2 YSLPELDYFSATEPYISGOINEIXY 27  
31 FTLPELDYDYSALPAVISPETIELHH 56

## RESULT 35

O82584 PRELIMINARY; PRT; 240 AA.

AC O82584; 01-NOV-1998 (TREMBlrel. 08, Created)  
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE Manganese superoxide dismutase (EC 1.15.1.1).  
GN MNSOD.  
OS Zantedeschia aethiopica (White calla lily).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Araceae; Zantedeschia.  
NCBI\_TaxId=69721;  
[1]

SEQUENCE FROM N.A.

RP TISSUE=LEAF;

RA Lino-Neto T., Tavares R.M., Palme K., Pais M.S.S.;

"Expression of superoxide dismutases during senescence and regreening of Zantedeschia aethiopica spathe.";

RT Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).

CC -1- CATALYTIC ACTIVITY: 2 SUPEROXIDE + 2 H(+) = O(2) + H(2)O(2).

CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE FAMILY.

DR EMBL: AF094832; AAC63379.1; -.  
DR HSSP: P04179; IABW.  
DR InterPro: IPR001189; SODismutase.  
DR Pfam: PF00081; sode; 1.  
DR Pfam: PF02777; sode; C; 1.  
DR ProDom: PD000475; SODismutase; 1.  
DR PROSITE: PS00088; SOD\_MN; 1.  
KM Oxidoreductase.

SQ SEQUENCE 240 AA; 26276 MW; 3D0479B8F38F4FDF CRC64;

Query Match 46.9%; Score 68; DB 10; Length 240;  
Best Local Similarity 46.2%; Pred. No. 0.023; Mismatches 7; Indels 0; Gaps 0;

QY 2 YSLPELDYFSATEPYISGOINEIXY 27  
Db 40 FSLPELDYDYSALPAVISPETIELHH 65

## RESULT 36

O9KM85 PRELIMINARY; PRT; 202 AA.

AC O9KM85; 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE Manganese superoxide dismutase.  
GN SODA.  
OS Vibrio vulnificus.

OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrrio.  
NCBI\_TaxId=672;  
[1]

SEQUENCE FROM N.A.

RC STRAIN=M2799;

RA Yamamoto S., Kimoto R.;

"Identification and characterization of the sda genes in some vibrio species.";

RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).

CC -1- CATALYTIC ACTIVITY: 2 SUPEROXIDE + 2 H(+) = O(2) + H(2)O(2).

CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE FAMILY.

DR EMBL: AB041847; BAA97807.1; -.  
DR HSSP: P09214; IMNG.

DR InterPro: IPR001189; SODismutase.

DR Pfam: PF00081; sode; 1.

DR Pfam: PF02777; sode; C; 1.

DR ProDom: PD000475; SODismutase; 1.

DR PROSITE: PS00088; SOD\_MN; 1.

KM Manganese; Oxidoreductase.

SQ SEQUENCE 202 AA; 22972 MW; 6177B5CF62CA700A CRC64;

Query Match 46.2%; Score 67; DB 2; Length 202;  
Best Local Similarity 40.7%; Pred. No. 0.027; Mismatches 10; Indels 0; Gaps 0;

QY 2 YSLPELDYFSATEPYISGOINEIXY 28  
Db 3 HTLPELDYDYSALPAVISPETIELHH 29

## RESULT 37

O9RUV2 PRELIMINARY; PRT; 211 AA.

AC O9RUV2; 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DT 01-JUN-2002 (T-EMBLrel. 21, last annotation update)  
 DE Superoxide dismutase (SOD), MN family.  
 GN DRI279.  
 OS Deinococcus radiodurans.  
 OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;  
 OC Deinococcaceae; Deinococcus.  
 NC NCB1\_TaxID=1299;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=RI;  
 RX MEDLINE=20036896; PubMed=10567266;  
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,  
 Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,  
 Vamathevan J., Lam P., McDonald L., Utterback T., Zaleski C.,  
 Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,  
 Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,  
 Fraser C.M.;  
 "Genome sequence of the radioresistant bacterium Deinococcus  
 radiodurans R1.";  
 RL Science 286:1571-1577(1999).  
 CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE  
 CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: 2 PEROXIDE RADICAL + 2 H(+) = O(2) + H(2)O(2).  
 CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE  
 FAMILY.  
 CC EMBL: AE001975; AAF10851.1; -.  
 DR HSSP: P00448; IVEW.  
 DR TIGR: DRI279; -.  
 DR InterPro: IPR001189; SODismutase.  
 DR Pfam: PF00081; sodfe, 1.  
 DR Pfam: PF02777; sodfe, C, 1.  
 DR ProDom: PD000475; SODismutase; 1.  
 DR PROSITE: PS00088; SOD MN; 1.  
 KM Oxidoreductase; Complete proteome.  
 SQ SEQUENCE 211 AA; 25478 MW; A71P22F516A1897E CRC64;  
 QY Query Match 46.2%; Score 67; DB 16; Length 211;  
 Best Local Similarity 44.4%; Pred. No. 0.029;  
 Matches 12; Conservative 6; Mismatches 9; Indels 0; Gaps 0;  
 QY 2 YSPELDYFSATEPYISQINEIXYT 28  
 Db 3 YTLPLDYALDALEPHIDARTMEIHT 29  
 SULT 38  
 ID Q9NB66 PRELIMINARY; PRT; 218 AA.  
 AC Q9NB66;  
 DT 01-OCT-2000 (T-EMBLrel. 15, Created)  
 DT 01-OCT-2000 (T-EMBLrel. 15, last sequence update)  
 DT 01-JUN-2002 (T-EMBLrel. 21, last annotation update)  
 DE Mitochondrial manganese superoxide dismutase precursor.  
 OS Callinectes sapidus (Blue crab).  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;  
 OC Eubrachyura; Portunoidae; Portunidae; Callinectes.  
 NC NCB1\_TaxID=6763;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Browner M., Browner T.H.;  
 RT "cDNA sequence of the mitochondrial manganese superoxide dismutase  
 from the blue crab, Callinectes sapidus.";  
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE  
 CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: 2 SUPEROXIDE + 2 H(+) = O(2) + H(2)O(2).  
 CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE  
 FAMILY.  
 CC EMBL: AF264029; AAF74770.1; -.  
 DR HSSP: P04179; IAP6.  
 DR InterPro: IPR001189; SODismutase.

DR Pfam: PF00081; sodfe, 1.  
 DR Pfam: PF02777; sodfe, C, 1.  
 DR ProDom: PD000475; SODismutase; 1.  
 DR PROSITE: PS00088; SOD MN; 1.  
 KM Oxidoreductase; Transferrin peptide.  
 FT TRANSIT 1 22  
 SQ SEQUENCE 218 AA; 24125 MW; A7FC617E8908CF0 CRC64;  
 QY Query Match 46.2%; Score 67; DB 5; Length 218;  
 Best Local Similarity 39.3%; Pred. No. 0.03;  
 Matches 11; Conservative 10; Mismatches 7; Indels 0; Gaps 0;  
 QY 1 YSPELDYFSATEPYISQINEIXYT 28  
 Db 23 KHTLPDLPYDGALEPISABIMQLHHS 50  
 RESULT 39  
 ID Q9SM64 PRELIMINARY; PRT; 228 AA.  
 AC Q9SM64;  
 DT 01-MAY-2000 (T-EMBLrel. 13, Created)  
 DT 01-MAY-2000 (T-EMBLrel. 13, last sequence update)  
 DT 01-JUN-2002 (T-EMBLrel. 21, last annotation update)  
 DE Manganese superoxide dismutase 1 precursor (BC 1.15.1.1).  
 GN SOD.  
 OS Prunus persica (Peach).  
 OC Mitochondrion.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eucosids 1; Rosales; Rosaceae; Amygdaloidae; Prunae.  
 OC NCB1\_TaxID=3760;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LEAF.  
 RA Bagnoli F., Racchi M.L.;  
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE  
 CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: 2 SUPEROXIDE + 2 H(+) = O(2) + H(2)O(2).  
 CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE  
 FAMILY.  
 CC EMBL: AJ238316; CAB56851.1; -.  
 DR HSSP: P04179; IAP6.  
 DR InterPro: IPR001189; SODismutase.  
 DR Pfam: PF00081; sodfe, 1.  
 DR Pfam: PF02777; sodfe, C, 1.  
 DR ProDom: PD000475; SODismutase; 1.  
 DR PROSITE: PS00088; SOD MN; 1.  
 KM Manganese; Mitochondrion; Oxidoreductase; Transferrin peptide.  
 FT TRANSIT 1 24  
 FT CHAIN 25 228  
 SQ SEQUENCE 228 AA; 25455 MW; A32DA57F552F18AB CRC64;  
 QY Query Match 46.2%; Score 67; DB 10; Length 228;  
 Best Local Similarity 46.2%; Pred. No. 0.031;  
 Matches 12; Conservative 6; Mismatches 8; Indels 0; Gaps 0;  
 QY 2 YSPELDYFSATEPYISQINEIXYT 27  
 Db 28 FSLPDLPPYDGALEPISGDIWQLHH 53  
 RESULT 40  
 ID Q43273 PRELIMINARY; PRT; 235 AA.  
 AC Q43273;  
 DT 01-NOV-1996 (T-EMBLrel. 01, Created)  
 DT 01-NOV-1996 (T-EMBLrel. 01, last sequence update)  
 DT 01-JUN-2002 (T-EMBLrel. 21, last annotation update)  
 DE Manganese superoxide dismutase (SOD-3).  
 OS Zea mays (Maize).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;  
 OC Panicoideae; Andropogoneae; Zea.  
 OX NCBI\_TaxID=4577;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89051020; PubMed=2461225;  
 RA White J.A., Scandalios J.G.;  
 RT "Isolation and characterization of a cDNA for mitochondrial manganese  
 superoxide dismutase (SOD-3) of maize and its relation to other  
 RT manganese superoxide dismutases.";  
 RL Biochim. Biophys. Acta 951:61-70(1988).  
 CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE  
 CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: 2 SUPEROXIDE + 2 H(+) = O(2) + H(2)O(2).  
 CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE  
 CC FAMILY.  
 CC DR EMBL; M33119; AAA33512.1; -.  
 DR HSSP; P04179; 1ABM.  
 DR InterPro; IPR001189; SODismutase.  
 DR Pfam; PF00081; sodfe; 1.  
 DR Pfam; PF02777; sodfe; 1.  
 DR ProDom; PD000475; SODismutase; 1.  
 DR PROSITE; PS00088; SOD\_MN; 1.  
 DR KX Oxidoreductase.  
 SQ SEQUENCE 235 AA; 25579 MW; AD5CDADD92FD336 CRC64;

Query Match 46.2%; Score 67; DB 10; Length 235;  
 Best Local Similarity 48.0%; Pred. No. 0.032; 7; Indels 0; Gaps 0;  
 Matches 12; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

Qy 3 SLPELDYFSATPEYISQINEIXY 27  
 Db 36 TLPLSYDFGALPAISGEIMRLH 60

Search completed: April 9, 2003, 13:59:51  
 Job time : 17.8624 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 9, 2003, 13:38:19 ; Search time 14.3119 Seconds

(without alignments)  
279.314 Million cell updates/sec

Title: US-09-987-190-2

Perfect score: 145  
Sequence: 1 KYSLEPDIYFSPSTPEYISQINEIXYTX 30

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

A\_Geneseq\_101002:\*

- 1: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:\*
- 2: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:\*
- 3: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:\*
- 4: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:\*
- 5: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:\*
- 6: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:\*
- 7: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:\*
- 8: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:\*
- 9: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:\*
- 10: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:\*
- 11: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:\*
- 12: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:\*
- 13: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:\*
- 14: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:\*
- 15: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:\*
- 16: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:\*
- 17: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:\*
- 18: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:\*
- 19: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:\*
- 20: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:\*
- 21: /SID52/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:\*
- 22: /SID52/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:\*
- 23: /SID52/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	141	97.2	30	19 AAM53252	Candida albicans a
2	100	69.0	233	23 ABB05612	S. cerevisiae mang
3	79	54.5	30	18 AAM29762	Malassezia fungus
4	79	54.5	224	18 AAM29771	Malassezia fungus
5	77	53.1	200	22 AAC92973	C glutamicum prote
6	77	53.1	200	22 AAU00514	C. melaleucola sup
7	74	51.0	101	21 AAG02172	Human secreted pro
8	74	51.0	155	22 AAU03130	Novel human secret
9	74	51.0	183	20 AA129656	Human manganese su
10	74	51.0	198	9 AAP80602	Sequence of polype

11	74	51.0	198	9 AAP80603	Sequence of polype
12	74	51.0	198	16 AAR75191	Human manganese su
13	74	51.0	198	16 AAR75192	Human manganese su
14	74	51.0	198	21 AAY55846	Human manganese su
15	74	51.0	199	13 AAR20015	Mn-SOD (1159Thr) .
16	74	51.0	199	15 AAR60359	MnSOD N-terminal.
17	74	51.0	222	8 AAP71701	Recombinant human
18	74	51.0	222	14 AAR44801	MnSOD cDNA. Homo
19	74	51.0	222	15 AAR61337	Human manganese-co
20	74	51.0	222	16 AAR75193	Human manganese su
21	74	51.0	222	17 AAW00018	Human manganese su
22	74	51.0	222	17 AAW00453	Human manganese su
23	74	51.0	222	17 AAR90713	Human manganese su
24	74	51.0	222	20 AAW98172	Human manganese su
25	74	51.0	222	20 AAW98173	Human manganese su
26	74	51.0	222	20 AAW98174	Human manganese su
27	74	51.0	222	20 AAW98175	Human manganese su
28	74	51.0	222	20 AAW98176	Human manganese su
29	74	51.0	222	20 AAW98177	Human manganese su
30	74	51.0	222	20 AAW98169	Human native manga
31	74	51.0	222	20 AAW98170	Human native manga
32	74	51.0	222	20 AAW98171	Human manganese su
33	74	51.0	222	20 AAW96317	Human manganese co
34	74	51.0	222	20 AAW82446	Human MnSOD protein
35	74	51.0	222	21 AAB08204	Amino acid sequenc
36	74	51.0	222	21 AAB08205	Q143N modified hum
37	74	51.0	222	21 AAB08206	A modified human m
38	74	51.0	222	23 ABB07330	Human manganese su
39	74	51.0	223	9 AAP80551	Human manganese su
40	74	51.0	226	21 AAY55849	Human manganese su
41	74	49.7	44	22 AAB78878	C. glutamicum SRT
42	71	49.0	188	19 AAM53253	Candida albicans f
43	71	49.0	207	19 AAB61476	A. fumigatus aller
44	70	48.3	46	9 AAP80604	Sequence of human
45	70	48.3	161	21 AAC33605	Arabidopsis thalia

#### ALIGNMENTS

RESULT 1	
AAM53252	
ID	AAM53252 standard; Protein; 30 AA.
XX	
AC	AAM53252;
XX	
DT	30-JUL-1998 (first entry)
XX	
DE	Candida albicans allergen - antigenic protein SEQ ID NO:2.
XX	
KW	Candida albicans; vaccine; allergen; antigenic protein; fungal antigen;
KW	immune response; infection; insoluble.
XX	
OS	Candida albicans.
XX	
FH	Key
FT	Misc-difference 26
FT	Misc-difference 29
FT	Misc-difference 30
FT	Misc-difference 30
FT	Misc-difference 30
XX	
PN	WO9809990-A1.
XX	
PD	12-MAR-1998.
XX	
PF	29-AUG-1997; 97MO-JP03041.
XX	
PR	31-MAR-1997; 97JP-0099775.
PR	04-SEP-1996; 96JP-0255400.
XX	
PA	(TAKI ) TAKARA SHUZO CO LTD.

XX Endo M, Kato I, Mizutani S, Takesako K;  
XX  
XX WPI; 1998-193553/17.  
DR  
XX Fungal antigens comprising insoluble fraction of fungal cells -  
PT useful for, e.g. stimulating immune response and treatment and  
PT diagnosis of fungal infection(s)  
XX  
PS Claim 26; Page 75; 108pp; Japanese.  
XX  
XX The present sequence represents an active vaccine component or allergen  
CC derived from *Candida albicans*, which is an antigenic protein. The  
CC present invention describes fungal antigens, comprising the insoluble  
CC fraction of fungal cells having completely/partially removed cell walls.  
CC Also described are nucleic acids encoding the antigens and a method for  
CC producing the antigens. The antigens can be used for preparing  
CC therapeutic compositions for stimulating immune response, e.g. as a  
CC vaccine. They can be used for treatment of fungal infections, treatment  
CC and prevention of allergies and diagnosis of fungal infections in  
CC vertebrates. The vaccines are not live, and have low toxicity.  
SQ  
SQ Sequence 30 AA;  
Query Match 97.2%; Score 141; DB 19; Length 30;  
Best Local Similarity 100.0%; Pred. No. 5.2e-16;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KYSLPELDYERFATEPYISGQINEIXYT 28  
DB 1 KYSLPELDYERFATEPYISGQINEIXYT 28  
RESULT 2  
ABB05612  
ID ABB05612 standard; Protein; 233 AA.  
XX  
AC ABB05612;  
XX  
DT 24-APR-2002 (first entry)  
XX  
DE S. cerevisiae manganese superoxide dismutase protein SEQ ID NO:31.  
XX  
KM *Aspergillus oryzae*; mutant; modified; reduced transcription; hormone;  
KM reduced translation; reduced secretion; receptor; antibody; reporter;  
KM enzyme; lipase; manganese superoxide dismutase.  
XX  
OS *Saccharomyces cerevisiae*.  
OS Synthetic.  
XX  
XX US6323002-B1.  
XX  
XX 27-NOV-2001.  
XX  
XX 25-JUN-1999; 99US-0339972.  
XX  
XX 12-SEP-1997; 97US-0928692.  
XX  
XX 13-SEP-1996; 96US-0713312.  
XX  
XX (NOVO) NOVOSYMS BIOTECH INC.  
XX  
XX Brody H, Yaver DS, Lamsa M, Hansen K;  
XX  
XX WPI; 2002-163017/21.  
DR  
XX Producing a polypeptide using a cell for reducing the production of the  
PT polypeptide, comprises inserting DNA into the genome of the cell at a  
PT position not within the polypeptide coding sequence or a regulatory  
PT sequence -  
XX  
XX Example 17; Column 125-128; 129pp; English.  
XX  
XX The present invention describes a method for producing a polypeptide (P1)  
CC

CC comprising cultivating a mutant cell whose parent cell comprises a DNA  
CC sequence encoding P1, by introducing a nucleic acid construct into the  
CC genome of the parent cell at a locus not within the P1 sequence, so  
CC that P1 transcription, translation or secretion is reduced, and  
CC recovering P1. The method is used to produce a polypeptide, such as a  
CC recombinant or heterologous hormone, hormone variant, receptor, antibody,  
CC reporter or enzyme, particularly an oxidoreductase, transferase,  
CC hydrolase, lyase, isomerase or ligase. The present sequence represents  
CC a *Saccharomyces cerevisiae* manganese superoxide dismutase protein which  
CC shares significant identity with the mutant *Aspergillus oryzae* DB9Y1058  
CC protein, which is used in an example from the present invention.  
XX  
SQ  
SQ Sequence 233 AA;  
Query Match 69.0%; Score 100; DB 23; Length 233;  
Best Local Similarity 64.3%; Pred. No. 3.8e-08;  
Matches 18; Conservative 5; Mismatches 5; Indels 0; Gaps 0;  
QY 1 KYSLPELDYERFATEPYISGQINEIXYT 28  
DB 27 KVTLPDLKMDFGALPEYISGQINEIXYT 54  
RESULT 3  
AAW29762  
ID AAW29762 standard; peptide; 30 AA.  
XX  
XX AAW29762;  
XX  
XX 20-FEB-1998 (first entry)  
XX  
XX Malassezia fungus MF-4 antigenic peptide.  
XX  
XX Malassezia; fungus; antigenic; human; IGE; immunoglobulin E;  
XX antibody; allergy; antigen.  
XX  
XX Malassezia sp.  
XX  
XX WO9721817-A1.  
XX  
XX 19-JUN-1997.  
XX  
XX 10-DEC-1996; 96WO-JP03602.  
XX  
XX 05-SEP-1996; 96JP-0257613.  
XX  
XX 12-DEC-1995; 95JP-0346627.  
XX  
XX 05-SEP-1996; 96JP-0257612.  
XX  
XX (TAKI) TAKARA SHUZO CO LTD.  
XX  
XX Akiyama K, Kato I, Kuroda M, Okado T, Onishi Y;  
XX  
XX Takesako K, Yagihara T, Yamaguchi H, Yasueda H;  
XX  
XX WPI; 1997-332788/30.  
XX  
XX Antigenic proteins from the fungus *Malassezia* - bind to IGE  
PT antibodies present in patients with *Malassezia* allergies, useful for  
PT diagnosis, treatment and prevention of such conditions  
XX  
XX Claim 8; Page 113; 162pp; Japanese.  
XX  
XX The present sequence represents a specifically claimed antigenic  
CC peptide isolated from the fungus *Malassezia*. The antigenic peptide  
CC can bind to IGE antibodies present in patients with allergic  
CC conditions. Antigenic proteins, peptides and nucleic acids from the  
CC fungus *Malassezia* can be used in the diagnosis, treatment and  
CC prevention of allergic conditions due to *Malassezia* organisms (such  
CC as *M. furfur*, *M. sympodialis* and *M. pachydermatitis*).  
XX  
SQ  
SQ Sequence 30 AA;  
Query Match 54.5%; Score 79; DB 18; Length 30;  
Best Local Similarity 55.6%; Pred. No. 8.4e-06;  
CC



XX 21-FEB-2001.  
PD  
XX  
PF 02-AUG-2000; 2000EP-0116669.  
XX  
PR 13-AUG-1999; 99US-0373731.  
XX  
PA (DEGS ) DEGUSSA-HUELS AG.  
XX (CNRS ) CNRS CENT NAT RECH SCI.  
PI Merkm M, Guyonvarch A, Marx A;  
XX WPI; 2001-212718/22.  
DR N-PSDB; AAS00520.  
XX  
PT New DNA encoding superoxide dismutase of *Corynebacterium*, useful for  
PT producing transformants with increased production of metabolites,  
PT particularly lysine -  
XX  
PS Claim 3; Page 14-15; 19pp; English.  
XX  
CC The sequence represents *Corynebacterium melassecola* superoxide dismutase  
CC (sod) polypeptide. *Corynebacterium* microorganisms may be transformed with a  
CC sod DNA sequence and the sod gene can be amplified. Sod is then often  
CC overexpressed in *Corynebacterium* bacteria. *Corynebacterium* that overexpress sod  
CC are used for production of metabolites, particularly nucleotides,  
CC vitamins and amino acids, especially D-pantoic acid or, specifically,  
CC L-lysine. These metabolites are useful in human or animal nutrition and  
CC as pharmaceuticals. Overexpression of sod in *Corynebacterium* increases the  
CC yield of particular metabolites.  
XX  
SQ Sequence 200 AA;  
XX  
Query March 53.1%; Score 77; DB 22; Length 200;  
Best Local Similarity 48.1%; Pred. No. 0.00019;  
Matches 13; Conservative 7; Mismatches 7; Indels 0; Gaps 0;  
QY 2 YSLPELDYFSAPEYISQINEIXYT 28  
DB 4 YELPELDYADYALPEPHIAEIMELHHS 30  
XX  
RESULT 7  
AAG02172  
ID AAG02172 standard; Protein; 101 AA.  
XX  
AC AAG02172;  
XX  
OT 06-OCT-2000 (first entry)  
XX  
E Human secreted protein, SEQ ID NO: 6253.  
XX  
E Human, 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
KW gene therapy; chromosome mapping.  
XX  
OS Homo sapiens.  
XX  
PN EP1033401-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 21-FEB-2000; 2000EP-0200610.  
XX  
PR 26-FEB-1999; 99US-0122487.  
XX  
PA (GEST ) GENSET.  
XX  
PI Dunas Milne Edwards J, Duclert A, Giordano J;  
XX WPI; 2000-500381/45.  
DR N-PSDB; AAC02178.  
XX  
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for

PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -  
XX  
PS Claim 13; SEQ ID 6253; 71pp + CD-ROM; English.  
XX  
CC The present sequence is a polypeptide encoded by one of a large number  
CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs  
CC were prepared from total human RNAs or polyA+ RNAs derived from 30  
CC different tissues. EST sequences usually correspond mainly to the 3'  
CC untranslated region (UTR) of the mRNA because they are often obtained  
CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for  
CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in  
CC those cases where longer cDNA sequences have been obtained, the full 5'  
CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'  
CC ends and can therefore be used to obtain full length cDNAs and genomic  
CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and  
CC chromosome mapping procedures. They are used to obtain upstream  
CC regulatory sequences and to design expression and secretion vectors.  
XX  
SQ Sequence 101 AA;  
XX  
Query March 51.0%; Score 74; DB 21; Length 101;  
Best Local Similarity 42.9%; Pred. No. 0.00026;  
Matches 13; Conservative 10; Mismatches 6; Indels 0; Gaps 0;  
QY 1 KYSLPELDYFSAPEYISQINEIXYT 28  
DB 25 KHSPLDLPDYDGALPEPHINQIMQLHHS 52  
XX  
RESULT 8  
AAU30130  
ID AAU30130 standard; Protein; 155 AA.  
XX  
AC AAU30130;  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Novel human secreted protein #621.  
XX  
KW Human; vaccination; gene therapy; nutritional supplement;  
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
XX  
OS Homo sapiens.  
XX  
PN WO200179449-A2.  
XX  
PD 25-OCT-2001.  
XX  
PF 16-APR-2001; 2001WO-US08656.  
XX  
PR 18-APR-2000; 2000US-0552929.  
XX 26-JAN-2001; 2001US-0770160.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Drmanac RT;  
XX WPI; 2001-611725/70.  
XX  
PT Nucleic acids encoding a range of human polypeptides, useful in genetic  
PT vaccination, testing and therapy -  
XX  
PS Claim 20; Page 248; 765pp; English.  
XX  
CC The invention relates to novel human secreted polypeptides. The  
CC polypeptides and antibodies to the polypeptides are useful for  
CC determining the presence of or predisposition to a disease associated  
CC with altered levels of polypeptide. The polypeptides are also useful for  
CC identifying agents (agonists and antagonists) that bind to them. Cells  
CC expressing the proteins are useful for identifying a therapeutic agent  
CC for use in treatment of a pathology related to aberrant expression or



CC physiological interactions of the polypeptide. Vectors comprising  
CC the nucleic acids encoding the polypeptides and cells genetically  
CC engineered to express them are also useful for producing the proteins.  
CC The proteins are useful in genetic vaccination, testing and  
CC therapy, and can be used as nutritional supplements. They may be used to  
CC increase stem cell proliferation, to regulate haematopoiesis, and in  
CC bone, cartilage, tendon, and/or nerve tissue growth or regeneration;  
CC immune suppression and/or stimulation; as anti-inflammatory agents; and  
CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid  
CC sequences of novel human secreted proteins of the invention.

XX Sequence 155 AA;

Query Match 51.0%; Score 74; DB 22; Length 155;  
Best Local Similarity 42.9%; Pred. No. 0.00044;  
Matches 12; Conservative 10; Mismatches 6; Indels 0; Gaps 0;

Qy 1 KYSLPDLPEFSATEPTISQGINEIXYT 28  
Db 50 KHSLPDLPYDGALEPHINQIMQLHHS 77

#### RESULT 9

AAZ29656  
ID AAY29656 standard; Protein; 183 AA.

XX AAY29656;

DT 03-NOV-1999 (first entry)

DE Human manganese superoxide dismutase exon 3-deleted isoform protein.

XX Human, manganese superoxide dismutase; MnSOD; exon 3 deleted isoform;

KM MnSOD E3(-); diagnosis; viral infection; HIV; oncogenesis; tumour;  
KM UV-induced damage; post ischemia reperfusion damage; anti-inflammatory;  
KM cytotoxic; cardiotoxic; cancer.

XX Homo sapiens.

OS Synthetic.

XX WO9943697-A1.

PD 02-SEP-1999.

PF 25-FEB-1999; 99WO-US04129.

PR 25-FEB-1998; 98US-0075948.

XX (UYCO-) UNIV & COMMUNITY COLLEGE.

XX Anziano PQ;

PI WPI; 1999-527592/44.

DR N-PSDB; AAZ08755.

XX A new isoform useful for diagnosing oxidative stress, and treating  
PT viral infections

PS Claim 1; Page 12; 64pp; English.

XX The present sequence represents an isoform of manganese superoxide  
CC dismutase (MnSOD), which is a splice-variant lacking exon 3 of the full  
CC length MnSOD (i.e. MnSOD E3(-)). MnSOD E3(-) may be useful in the  
CC treatment of a wide variety of disorders including viral infections,  
CC particularly HIV, and may be used for the prevention of oncogenesis,  
CC tumour promotion and invasiveness, and UV-induced damage, for protection  
CC of cardiac tissue against post ischemia reperfusion damage, as an  
CC anti-inflammatory agent, to reduce the cytotoxic and cardiotoxic effects  
CC of anti-cancer drugs, and to improve the longevity of living cells.

XX Sequence 183 AA;

Query Match 51.0%; Score 74; DB 20; Length 183;

Best Local Similarity 42.9%; Pred. No. 0.00054;  
Matches 12; Conservative 10; Mismatches 6; Indels 0; Gaps 0;

Qy 1 KYSLPDLPEFSATEPTISQGINEIXYT 28  
Db 25 KHSLPDLPYDGALEPHINQIMQLHHS 52

#### RESULT 10

AAAP80602  
ID AAP80602 standard; Protein; 198 AA.

XX AAP80602;

DT 03-OCT-1990 (first entry)

DE Sequence of polypeptide of Formula IVA encoding modified portion of  
DE human manganese superoxide dismutase (MnSOD).

XX Human manganese superoxide dismutase derivative; clone BS8;  
KM formula IVA; enzyme; EC-1.15.1.1.

XX Homo sapiens.

PN EP282899-A.

PD 21-SEP-1988.

PF 10-MAR-1988; 88EP-0103754.

PR 24-DEC-1987; 87DE-3744038.

PR 14-MAR-1987; 87DE-3708306.

XX (BOEH) BOEHRINGER INELHEIM.

XX Heckl K, Spevak W, Ostermann E, Zophel A, Krystek E, Maurer-Fogy I;  
PI White-Castano MJ, Stratowa C, Hauptmann R;

DR WPI; 1988-265361/38.

DR N-PSDB; AAN61218.

XX Prodn. of human manganese superoxidisedismutase peptide(s) -  
PT and DNA coding sequences, for control and diagnosis of eg  
PT inflammatory diseases

PS Disclosure; Page 8; 57pp; German.

XX The patent is for polypeptides, esp. non-glycosylated, having the  
CC enzymatic, biological and immunological properties of MnSOD that are  
CC prep. by genetic engineering methods. Also new are DNA sequences  
CC encoding all or part of the polypeptides and replicating vectors,  
CC expression plasmids and transformed host cells contg. these sequences.  
CC MnSOD catalyses disproportionation of the superoxide radical and can be  
CC used in the prevention, diagnosis and treatment of inflammatory,  
CC degenerative, neoplastic and rheumatoid disorders; for wound healing; in  
CC autoimmune disease and in organ transplantation. It can also be used to  
CC improve the storage stability of liq. and solid foods.

XX Sequence 198 AA;

Query Match 51.0%; Score 74; DB 9; Length 198;  
Best Local Similarity 42.9%; Pred. No. 0.0006;  
Matches 12; Conservative 10; Mismatches 6; Indels 0; Gaps 0;

Qy 1 KYSLPDLPEFSATEPTISQGINEIXYT 28  
Db 1 KHSLPDLPYDGALEPHINQIMQLHHS 28

#### RESULT 11

AAAP80603  
ID AAP80603 standard; Protein; 198 AA.

XX AAP80603



Key	Location/Qualifiers
Modified-site	39..41 /label= N-glycosylation_site /note= "potential"
Modified-site	73..75 /label= N-glycosylation_site /note= "potential"
Modified-site	80..82 /label= N-glycosylation_site /note= "potential"
Modified-site	188..190 /label= N-glycosylation_site /note= "potential"
EP676472-A1.	
11-OCT-1995.	
10-MAR-1988;	88EP-0107460.
24-DEC-1987;	87DE-3744038.
14-MAY-1987;	87DE-3708306.
26-MAY-1987;	87DE-3717695.
10-JUL-1987;	87DE-3722884.
(BOEH ) BOEHRINGER INGELHEIM INT GMBH.	
Hauptmann R, Heckl K, Krystek E, Maurer-Fogy I, Ostermann E, Spevak W, Stratowa C, Wache-Castanon M, Zoepfel A;	
WPI; 1995-346092/45.	
N-PSDB; AAQ94272.	
Genes encoding recombinant human manganese super:oxide dismutase - for treatment, prevention and diagnosis of inflammatory diseases	
Claim 1; Page 9-10; 54pp; German.	
The present sequence is that of the 198 amino acid long hMn-SOD mature polypeptide having a Gln residue at position 29. The protein is produced with an N-terminal Met residue which is then removed, e.g. by treatment with CNBr or CNCl, to release mature hMn-SOD. The mature recombinant protein is useful for treating, preventing or diagnosing inflammatory, degenerative, neoplastic and rheumatic disorders; in wound healing and in autoimmune diseases.	
Sequence 198 AA;	
Query Match	51.0%; Score 74; DB 16; Length 198;
Best Local Similarity	42.9%; Pred. No. 0.0006;
Matches 12; Conservative 10; Mismatches 6; Indels 0; Gaps 0;	
1 KXSLPELDYBSATPEYISGQINEIXYT 28	
1 KXSLPDLPYDYGALPEPHINAQIMOLHNS 28	
RESULT 14	
AAV55846	
ID AAV55846 standard; Protein; 198 AA.	
AAV55846;	
06-MAR-2000 (first entry)	
Human manganese superoxide dismutase (MnSOD) protein.	
Manganese superoxide dismutase; MnSOD; EC:SOD; proteoglycan; human; endothelial; oxidative damage; pulmonary inflammatory injury; cancer; lung disease; hypoxia; ischemia; reperfusion injury; arthritis; hyperoxia; atherosclerosis; lupus erythematosus; hypertension;	

[illegible]



OY 1 KYSLPELDYEFSAPEYISGQINEIXYT 28  
 DB 25 KHSPLDLPYDYGALPEPHINAQIMQLHHS 52

## RESULT 18

AAR44801  
 ID AAR44801 standard; Protein; 222 AA.

AC AAR44801;

DT 21-JUN-1994 (first entry)

DE MnSOD cDNA.

MnSOD; manganese superoxide dismutase; N-terminal; catalyst;  
 KM reperfusion; injury; ischaemia; superoxide; SO; molecular oxygen;  
 KM anti-inflammatory; ss.

OS Homo sapiens.

PN USS270195-A.

PD 14-DEC-1993.

PF 22-NOV-1985; 85US-0801090.

PR 29-OCT-1986; 86IE-0002851.

PA (BIOT-) BIO-TECHNOLOGY GENERAL CORP.

PI Beck Y, Hartman JR;

PS Claim 1; Fig 1a-1c; 27pp; English.

DR WPI; 1993-404931/50.

DR N-PSDB; AAQ53193.

XX Expression plasmid in Escherichia coli host system - encodes

PT human manganese superoxidisedismutase analogue, useful for e.g.

PT treating inflammation

XX The sequence is of a manganese superoxide dismutase which can be used

CC to catalyse the reduction of superoxide (SO) radicals to hydrogen

CC peroxide and molecular oxygen. It can be used to reduce

CC reperfusion injury following ischaemia and prolong the survival of

CC excised organs. It can also be used as a long acting

CC anti-inflammatory drug.

XX Sequence 222 AA;

Query Match 51.0%; Score 74; DB 14; Length 222;

Best Local Similarity 42.9%; Pred. No. 0.00069;

Matches 12; Conservative 10; Mismatches 6; Indels 0; Gaps 0;

OY 1 KYSLPELDYEFSAPEYISGQINEIXYT 28

DB 25 KHSPLDLPYDYGALPEPHINAQIMQLHHS 52

RESULT 19

AAR61337

ID AAR61337 standard; Protein; 222 AA.

AC AAR61337;

DT 31-MAY-1995 (first entry)

DE Human manganese-contg. superoxide dismutase SOD-2.

KM Human superoxide dismutase; hSOD2; mSOD; neurodegeneration;

KW Alzheimer's disease; Parkinson's disease; Huntington's disease;

KM Hallervorden-Spatz disease; olivopontocerebellar atrophy;  
 KM familial amyotrophic lateral sclerosis; FALS; diagnosis; mutant SOD;  
 KM reverse transcription polymerase chain reaction; SSCP analysis.

OS Homo sapiens.

Key Location/Qualifiers

FT Peptide 1..18

FT Protein /label= signal peptide

FT 19..222

FT /label= mature\_mSOD

PN M09419493-A.

PD 01-SEP-1994.

PF 28-FEB-1994; 94MO-US02089.

PR 26-FEB-1993; 93US-0023980.

PA (GENO ) GEN HOSPITAL CORP.

PI (MASI ) MASSACHUSETTS INST TECHNOLOGY.

PI Brown R, Horvitz HR, Rosen DR;

PS WPI; 1994-294353/36.

DR N-PSDB; AAQ57474.

XX Diagnosis, treatment and prevention of diseases of cell death -

PT e.g. amyotrophic lateral sclerosis, which are the result of e.g.

PT decreased SOD activity

XX Claim 22; Fig 3; 94pp; English.

PS The presence of a mutation in a gene encoding a superoxide

CC dismutase (SOD1, SOD2 or SOD3) indicates an increased likelihood of

CC developing a cell death disease, specifically a neurodegenerative

CC disease. The use of SOD polypeptides to treat amyotrophic lateral

CC sclerosis and diseases involving a deleterious mutation in the

CC glutathione peroxidase-encoding gene, the catalase-encoding gene

CC and the nitric oxide-encoding gene is claimed. The specification

CC includes the sequences of human SOD1, hSOD2 and hSOD3 (AAR61336-

CC AAR61338, respectively).

XX Sequence 222 AA;

Query Match 51.0%; Score 74; DB 15; Length 222;

Best Local Similarity 42.9%; Pred. No. 0.00069;

Matches 12; Conservative 10; Mismatches 6; Indels 0; Gaps 0;

OY 1 KYSLPELDYEFSAPEYISGQINEIXYT 28

DB 25 KHSPLDLPYDYGALPEPHINAQIMQLHHS 52

RESULT 20

AAR75193

ID AAR75193 standard; Protein; 222 AA.

AC AAR75193;

DT 23-MAY-1996 (first entry)

DE Human manganese superoxide dismutase (lys29) preproteoin.

KM hMn-SOD; manganese; superoxide dismutase; tetramer; inflammation;

KW autoimmune disease; rheumatoid arthritis; neoplasm; emphysema;

KM osteoarthritis; wound healing; leader peptide; precursor.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..24



CC Human manganese superoxide dismutase (AAW00453) catalyses the  
 CC decomposition of the superoxide anion to peroxide. Agents such as  
 CC superoxide dismutase are able to increase the salvage of heart muscle  
 CC cells. The SOD gene (AAT15918) can be used in novel therapeutic  
 CC constructs also contg. a tissue-specific promoter and a hypoxia  
 CC response enhancer element. Expression of the gene is enhanced in the  
 CC target tissue under hypoxic conditions, such as those encountered  
 CC during ischaemia and reperfusion.

XX  
 SQ Sequence 222 AA;  
 Query Match 51.0%; Score 74; DB 17; Length 222;  
 Best Local Similarity 42.9%; Pred. No. 0.00069;  
 Matches 12; Conservative 10; Mismatches 6; Indels 0; Gaps 0;

OY 1 KXSLPELDYERFATPEYISQGINEIXYT 28  
 Db 25 KXSLPELDYERFATPEYISQGINEIXYT 52

RESULT 23  
 AAR90713  
 ID AAR90713 standard; Protein; 222 AA.  
 AC AAR90713;  
 XX  
 DT 06-APR-1996 (first entry)  
 XX  
 DE Human manganese superoxide-dismutase.  
 KW Human; manganese superoxide-dismutase; T-lymphocyte; probe; cDNA;  
 KW Escherichia coli; plasmid pMS-4; N-terminal truncation; cloning;  
 KM aminopeptidase; antiinflammatory; oxygen free radical scavenger;  
 KM synovial inflammation; arthritis; lung fibrosis.

XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..24  
 FT /note= "Pre-peptide"  
 FT Region 39..48  
 FT /note= "Region used to construct probe AAT15591"  
 FT Misc-difference 66  
 FT /note= "Gln in previously published sequence"  
 FT Misc-difference 112  
 FT /note= "Gln in previously published sequence"  
 FT Misc-difference 132  
 FT /note= "Gln in previously published sequence"  
 FT Misc-difference 147..148  
 FT /note= "Residues not present in previously published sequence"  
 FT Region 203..213  
 FT /note= "Region used to construct probe AAT15592"

XX  
 PN EP691401-A1.  
 PD 10-JAN-1996.  
 XX  
 PF 25-MAR-1988; 88EP-0106995.  
 XX  
 PR 26-FEB-1988; 88US-0161117.  
 PR 27-MAR-1987; 87US-0032734.  
 XX  
 PA (BIOT-) BIO-TECHNOLOGY GENERAL CORP.  
 PI Beck Y, Hartman JR, Nimrod A;  
 DR WPI; 1996-059735/07.  
 DR N-PSDB; AAT15589.  
 XX  
 PT New human manganese superoxide dismutase analogues - having a Lys  
 PT and Opt. His residue absent from the N-terminus, useful for treating  
 PT e.g. synovial inflammation, arthritis or lung fibrosis

XX  
 PS Example 1; Fig 1; 45pp; English.  
 XX  
 CC The sequence represents a human manganese superoxide-dismutase  
 CC (MnSOD), and is encoded by a human T-lymphocyte cDNA isolated from  
 CC a library in phage lambda-gt10 in Escherichia coli, by screening  
 CC with a 5'-probe (AAT15591) and a 3'-probe (AAT1592), followed by  
 CC subcloning in plasmid pBR322 to give plasmid pMS-4. The cDNA may  
 CC be expressed in e.g. E. coli for production of recombinant MnSOD.  
 CC The MnSOD product may be cleaved with Aeromonas proteolytica  
 CC aminopeptidase to produce an N-terminally truncated analogue with  
 CC lysine and optionally histidine residues removed. The MnSOD  
 CC analogue may be used in therapy of conditions associated with  
 CC generation of oxygen free radicals, particularly synovial  
 CC inflammation, arthritis and lung fibrosis (claimed).

XX  
 SQ Sequence 222 AA;  
 Query Match 51.0%; Score 74; DB 17; Length 222;  
 Best Local Similarity 42.9%; Pred. No. 0.00069;  
 Matches 12; Conservative 10; Mismatches 6; Indels 0; Gaps 0;

OY 1 KXSLPELDYERFATPEYISQGINEIXYT 28  
 Db 25 KXSLPELDYERFATPEYISQGINEIXYT 52

RESULT 24  
 AAM98172  
 ID AAM98172 standard; Protein; 222 AA.  
 AC AAM98172;  
 XX  
 DT 05-JUL-1999 (first entry)  
 XX  
 DE Human manganese superoxide dismutase Q143A mutant.  
 KW Superoxide dismutase; MnSOD; SOD; human; protein engineering;  
 KW enzyme engineering; antioxidant; antiinflammatory; cytotoxicity;  
 KM inflammation; reperfusion injury; therapy; mutant.

XX  
 OS Homo sapiens.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..24  
 FT /note= "signal peptide"  
 FT Protein 25..222  
 FT /note= "mature protein"  
 FT Active-site 50  
 FT /note= "His-26"  
 FT Active-site 54  
 FT /note= "His-30"  
 FT Active-site 58  
 FT /note= "Tyr-58"  
 FT Active-site 98  
 FT /note= "Tyr-74"  
 FT Active-site 167  
 FT /note= "Gln-143"  
 FT Active-site 183  
 FT /note= "Asp-159"  
 FT Active-site 185  
 FT /note= "Trp-161"  
 FT Active-site 187  
 FT /note= "His-163"  
 FT Misc-difference 155  
 FT /label= Gln, Gln  
 FT Misc-difference 167  
 FT /note= "Gln replaces Ala in wild-type hMnSOD"

XX  
 PN WO9913088-A1.  
 PN 18-MAR-1999.

XX 10-SEP-1998; 98WO-US18842.  
 XX 10-SEP-1997; 97US-0927230.  
 XX (UYFL ) UNIV FLORIDA.  
 XX Nick HS, Silverman DN;  
 XX WPI; 1999-229242/19.  
 FT New recombinant human manganese superoxide dismutase proteins  
 PS Claim 14; Page -; 61pp; English.  
 CC This protein is a Q143A mutant of human manganese superoxide  
 CC dismutase (hMnSOD, see also AAW98170). Novel modified hMnSOD  
 CC proteins having catalytic activity which differs from natural  
 CC hMnSOD are claimed. The modified proteins (see also AAW98171-77)  
 CC exhibit reduced or no product inhibition, or have greater  
 CC activity, or both, compared to natural hMnSOD. The modifications  
 CC involve one or amino acid substitutions within the active site of  
 CC the enzyme, especially at residues His-26, His-30, Tyr-34, His-74,  
 CC Gln-143, Asp-159, Trp-161 and His-163. Gln-143, Tyr-34 and His-30  
 CC are located near the active site metal, are highly conserved, and  
 CC are involved in a hydrogen bond relay which links the active site  
 CC metal-bound hydroxyl group to ordered solvent, and are thus all  
 CC potentially involved in catalytic activity, product inhibition  
 CC and/or enzyme stability. Modified hMnSODs, or expression vectors  
 CC in which a modified hMnSOD nucleic acid is linked to a promoter  
 CC (preferably mammalian), are used to protect a cell line from damage  
 CC caused by superoxide radicals (claimed). They can also be used to  
 CC treat subjects suffering from, or at risk of, cytotoxicity caused  
 CC by superoxide radicals (claimed). As such, they can be used as  
 CC antioxidants in the treatment of a variety of disorders, including  
 CC inflammation (claimed), reperfusion injury following ischemia  
 CC (claimed), and cellular damage caused by chemotherapeutic agents.  
 CC Note. The present sequence is not shown in the specification but  
 CC is derived from the hMnSOD sequence given on page 53-54.  
 XX  
 SQ Sequence 222 AA;  
 Query Match 51.0%; Score 74; DB 20; Length 222;  
 Best Local Similarity 42.9%; Pred. No. 0.00069;  
 Matches 12; Conservative 10; Mismatches 6; Indels 0; Gaps 0;  
 QY 1 KYSLPELDYFSATPEYISGQINEXYT 28  
 DB 25 KHSLPDLPYDYGALPEHINAQIMQIHHS 52  
 RESULT 25  
 AAW98173  
 ID AAW98173 standard; Protein; 222 AA.  
 XX  
 AC AAW98173;  
 XX  
 DT 05-JUL-1999 (first entry)  
 XX  
 DE Human manganese superoxide dismutase Q143H mutant.  
 XX  
 KW Superoxide dismutase; MnSOD; SOD; human; protein engineering;  
 KW enzyme engineering; antioxidant; antiinflammatory; cytotoxicity;  
 KW inflammation; reperfusion injury; therapy; mutant.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 FH Key  
 FH Peptide  
 FT Location/Qualifiers  
 FT 1..24  
 FT /note= "signal peptide"  
 FT 25..222  
 FT /note= "mature protein"

FT Active-site 50  
 FT /note= "His-26"  
 FT Active-site 54  
 FT /note= "His-30"  
 FT Active-site 58  
 FT /note= "Tyr-58"  
 FT Active-site 98  
 FT /note= "Tyr-74"  
 FT Active-site 167  
 FT /note= "Gln-143"  
 FT Active-site 183  
 FT /note= "Asp-159"  
 FT Active-site 185  
 FT /note= "Trp-161"  
 FT Active-site 187  
 FT /note= "His-163"  
 FT Misc-difference 155  
 FT /label= Gln, Glu  
 FT Misc-difference 167  
 FT /note= "Gln replaces His in wild-type hMnSOD"  
 XX  
 PN W09913088-A1.  
 XX  
 PD 18-MAR-1999.  
 XX  
 XX 10-SEP-1998; 98WO-US18842.  
 PF 10-SEP-1997; 97US-0927230.  
 XX  
 XX (UYFL ) UNIV FLORIDA.  
 XX Nick HS, Silverman DN;  
 XX WPI; 1999-229242/19.  
 DR  
 XX New recombinant human manganese superoxide dismutase proteins  
 FT  
 PS Claim 14; Page -; 61pp; English.  
 CC This protein is a Q143H mutant of human manganese superoxide  
 CC dismutase (hMnSOD, see also AAW98170). Novel modified hMnSOD  
 CC proteins having catalytic activity which differs from natural  
 CC hMnSOD are claimed. The modified proteins (see also AAW98171-77)  
 CC exhibit reduced or no product inhibition, or have greater  
 CC activity, or both, compared to natural hMnSOD. The modifications  
 CC involve one or amino acid substitutions within the active site of  
 CC the enzyme, especially at residues His-26, His-30, Tyr-34, His-74,  
 CC Gln-143, Asp-159, Trp-161 and His-163. Gln-143, Tyr-34 and His-30  
 CC are located near the active site metal, are highly conserved, and  
 CC are involved in a hydrogen bond relay which links the active site  
 CC metal-bound hydroxyl group to ordered solvent, and are thus all  
 CC potentially involved in catalytic activity, product inhibition  
 CC and/or enzyme stability. Modified hMnSODs, or expression vectors  
 CC in which a modified hMnSOD nucleic acid is linked to a promoter  
 CC (preferably mammalian), are used to protect a cell line from damage  
 CC caused by superoxide radicals (claimed). They can also be used to  
 CC treat subjects suffering from, or at risk of, cytotoxicity caused  
 CC by superoxide radicals (claimed). As such, they can be used as  
 CC antioxidants in the treatment of a variety of disorders, including  
 CC inflammation (claimed), reperfusion injury following ischemia  
 CC (claimed), and cellular damage caused by chemotherapeutic agents.  
 CC Note. The present sequence is not shown in the specification but  
 CC is derived from the hMnSOD sequence given on page 53-54.  
 XX  
 SQ Sequence 222 AA;  
 Query Match 51.0%; Score 74; DB 20; Length 222;  
 Best Local Similarity 42.9%; Pred. No. 0.00069;  
 Matches 12; Conservative 10; Mismatches 6; Indels 0; Gaps 0;  
 QY 1 KYSLPELDYFSATPEYISGQINEXYT 28  
 DB 25 KHSLPDLPYDYGALPEHINAQIMQIHHS 52



RESULT 26  
AAW98174  
ID AAW98174 standard; Protein: 222 AA.  
XX  
AC AAW98174;  
XX  
DT 05-JUL-1999 (first entry)  
XX  
DE Human manganese superoxide dismutase Q143D mutant.  
XX  
KM Superoxide dismutase; MnSOD; SOD; human; protein engineering;  
KM enzyme engineering; antioxidant; antiinflammatory; cytotoxicity;  
KM inflammation; reperfusion injury; therapy; mutant.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
XX Key Location/Qualifiers  
XX Peptide 1..24  
FT /note= "signal peptide"  
FT Protein 25..222  
FT /note= "mature protein"  
FT Active-site 50  
FT /note= "His-26"  
FT Active-site 54  
FT /note= "His-30"  
FT Active-site 58  
FT /note= "Tyr-58"  
FT Active-site 98  
FT /note= "Tyr-74"  
FT Active-site 167  
FT /note= "Gln-143"  
FT Active-site 183  
FT /note= "Asp-159"  
FT Active-site 185  
FT /note= "Tyr-161"  
FT Active-site 187  
FT /note= "His-163"  
FT Misc-difference 155  
FT /label= Gln, Glu  
FT Misc-difference 167  
FT /note= "Gln replaces Asp in wild-type hMnSOD"  
XX  
XX W09913088-A1.  
XX  
XX 18-MAR-1999.  
XX  
XX 10-SEP-1998; 98MO-US18842.  
XX  
XX 10-SEP-1997; 97US-0927230.  
XX  
XX (UYFL ) UNIV FLORIDA.  
XX  
XX Nick HS, Silverman DN;  
XX  
XX WPI; 1999-229242/19.  
XX  
XX New recombinant human manganese superoxide dismutase proteins  
XX  
XX Claim 14; Page -: 61pp; English.  
XX  
XX This protein is a Q143D mutant of human manganese superoxide  
XX dismutase (hMnSOD, see also AAW98170). Novel modified hMnSOD  
XX proteins having catalytic activity which differs from natural  
XX hMnSOD are claimed. The modified proteins (see also AAW98171-77)  
XX exhibit reduced or no product inhibition, or have greater  
XX activity, or both, compared to natural hMnSOD. The modifications  
XX involve one or amino acid substitutions within the active site of  
XX the enzyme, especially at residues His-26, His-30, Tyr-34, His-74,  
XX Gln-143, Asp-159, Tyr-161 and His-163. Gln-143, Tyr-34 and His-30  
XX are located near the active site metal, are highly conserved, and

CC are involved in a hydrogen bond relay which links the active site  
CC metal-bound hydroxyl group to ordered solvent, and are thus all  
CC potentially involved in catalytic activity, product inhibition  
CC and/or enzyme stability. Modified hMnSODs, or expression vectors  
CC in which a modified hMnSOD nucleic acid is linked to a promoter  
CC (preferably mammalian), are used to protect a cell line from damage  
CC caused by superoxide radicals (claimed). They can also be used to  
CC treat subjects suffering from, or at risk of, cytotoxicity caused  
CC by superoxide radicals (claimed). As such, they can be used as  
CC antioxidants in the treatment of a variety of disorders, including  
CC inflammation (claimed), reperfusion injury following ischemia  
CC (claimed), and cellular damage caused by chemotherapeutic agents.  
CC Note. The present sequence is not shown in the specification but  
CC is derived from the hMnSOD sequence given on page 53-54.  
XX  
SQ Sequence 222 AA;  
Query Match 51.0%; Score 74; DB 20; Length 222;  
Best Local Similarity 42.9%; Pred. No. 0.00069;  
Matches 12; Conservative 10; Mismatches 6; Indels 0; Gaps 0;  
QY 1 KYSLPELDYERFATEPIYSGQINEIXYT 28  
Db 25 KHSLPDLPYDYGALDEPHINQIMQIHHS 52  
RESULT 27  
AAW98175  
ID AAW98175 standard; Protein: 222 AA.  
XX  
AC AAW98175;  
XX  
DT 05-JUL-1999 (first entry)  
XX  
DE Human manganese superoxide dismutase Q143E mutant.  
XX  
XX  
XX Superoxide dismutase; MnSOD; SOD; human; protein engineering;  
XX enzyme engineering; antioxidant; antiinflammatory; cytotoxicity;  
KM inflammation; reperfusion injury; therapy; mutant.  
XX  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
XX Key Location/Qualifiers  
XX Peptide 1..24  
FT /note= "signal peptide"  
FT Protein 25..222  
FT /note= "mature protein"  
FT Active-site 50  
FT /note= "His-26"  
FT Active-site 54  
FT /note= "His-30"  
FT Active-site 58  
FT /note= "Tyr-58"  
FT Active-site 98  
FT /note= "Tyr-74"  
FT Active-site 167  
FT /note= "Gln-143"  
FT Active-site 183  
FT /note= "Asp-159"  
FT Active-site 185  
FT /note= "Tyr-161"  
FT Active-site 187  
FT /note= "His-163"  
FT Misc-difference 155  
FT /label= Gln, Glu  
FT Misc-difference 167  
FT /note= "Gln replaces Glu in wild-type hMnSOD"  
XX  
XX W09913088-A1.  
XX  
XX 18-MAR-1999.

PF 10-SEP-1998; 98WO-US18842.  
 XX  
 PR 10-SEP-1997; 97US-0927230.  
 XX (UYFL ) UNIV FLORIDA.  
 PA  
 XX Nick HS, Silverman DN;  
 XX WPI; 1999-229242/19.  
 DR  
 XX  
 FT New recombinant human manganese superoxide dismutase proteins  
 PS  
 XX Claim 14; Page -; 61pp; English.  
 CC This protein is a Q143E mutant of human manganese superoxide  
 CC dismutase (hMnSOD, see also AAW98170). Novel modified hMnSOD  
 CC proteins having catalytic activity which differs from natural  
 CC hMnSOD are claimed. The modified proteins (see also AAW98171-77)  
 CC exhibit reduced or no product inhibition, or have greater  
 CC activity, or both, compared to natural hMnSOD. The modifications  
 CC involve one or amino acid substitutions within the active site of  
 CC the enzyme, especially at residues His-26, His-30, Tyr-34, His-74,  
 CC Gln-143, Asp-159, Trp-161 and His-163. Gln-143, Tyr-34 and His-30  
 CC are located near the active site metal, are highly conserved, and  
 CC are involved in a hydrogen bond relay which links the active site  
 CC metal-bound hydroxyl group to ordered solvent, and are thus all  
 CC potentially involved in catalytic activity, product inhibition  
 CC and/or enzyme stability. Modified hMnSODs, or expression vectors  
 CC in which a modified hMnSOD nucleic acid is linked to a promoter  
 CC (preferably mammalian), are used to protect a cell line from damage  
 CC caused by superoxide radicals (claimed). They can also be used to  
 CC treat subjects suffering from, or at risk of, cytotoxicity caused  
 CC by superoxide radicals (claimed). As such, they can be used as  
 CC antioxidants in the treatment of a variety of disorders, including  
 CC inflammation (claimed), reperfusion injury following ischemia  
 CC (claimed), and cellular damage caused by chemotherapeutic agents.  
 CC Note. The present sequence is not shown in the specification but  
 CC is derived from the hMnSOD sequence given on page 53-54.  
 XX  
 SO Sequence 222 AA;  
 Query Match 51.0%; Score 74; DB 20; Length 222;  
 Best Local Similarity 42.9%; Pred. No. 0.00069;  
 Matches 12; Conservative 10; Mismatches 6; Indels 0; Gaps 0;  
 Oy 1 KYSLPELDYFSATEPYISGOINEXYT 28  
 Db 25 KHSPLDLPYDYGALPHINAQIMQLHHS 52  
 RESULT 28  
 ID AAW98176 standard; Protein; 222 AA.  
 XX  
 AC AAW98176;  
 XX  
 DT 05-JUL-1999 (first entry)  
 XX  
 DE Human manganese superoxide dismutase H30N mutant.  
 XX  
 KM Superoxide dismutase; MnSOD; SOD; human; protein engineering;  
 KM enzyme engineering; antioxidant; antiinflammatory; cytotoxicity;  
 KM inflammation; reperfusion injury; therapy; mutant.  
 OS Homo sapiens.  
 OS Synthetic.  
 OS  
 FH Key Location/Qualifiers  
 FT Peptide 1..24  
 FT Protein /note= "signal peptide"  
 FT 25..222  
 FT /note= "mature protein"  
 FT Active-site 50

FT /note= "His-26"  
 FT Active-site 54  
 FT /note= "His-30"  
 FT Active-site 58  
 FT /note= "Tyr-58"  
 FT Active-site 98  
 FT /note= "Tyr-74"  
 FT Active-site 167  
 FT /note= "Gln-143"  
 FT Active-site 183  
 FT /note= "Asp-159"  
 FT Active-site 185  
 FT /note= "Trp-161"  
 FT Active-site 187  
 FT /note= "His-163"  
 FT Misc-difference 155  
 FT /label= Gln, Glu  
 FT Misc-difference 54  
 FT /note= "His replaces Asn in wild-type hMnSOD"  
 FT  
 FT W09913086-A1.  
 PD 18-MAR-1999.  
 XX  
 XX 10-SEP-1998; 98WO-US18842.  
 PF 10-SEP-1997; 97US-0927230.  
 XX (UYFL ) UNIV FLORIDA.  
 XX  
 PA Nick HS, Silverman DN;  
 XX WPI; 1999-229242/19.  
 DR  
 XX  
 FT New recombinant human manganese superoxide dismutase proteins  
 PS  
 XX Claim 33; Page -; 61pp; English.  
 CC This protein is a H30N mutant of human manganese superoxide  
 CC dismutase (hMnSOD, see also AAW98170). Novel modified hMnSOD  
 CC proteins having catalytic activity which differs from natural  
 CC hMnSOD are claimed. The modified proteins (see also AAW98172-77)  
 CC exhibit reduced or no product inhibition, or have greater  
 CC activity, or both, compared to natural hMnSOD. The modifications  
 CC involve one or amino acid substitutions within the active site of  
 CC the enzyme, especially at residues His-26, His-30, Tyr-34, His-74,  
 CC Gln-143, Asp-159, Trp-161 and His-163. Gln-143, Tyr-34 and His-30  
 CC are located near the active site metal, are highly conserved, and  
 CC are involved in a hydrogen bond relay which links the active site  
 CC metal-bound hydroxyl group to ordered solvent, and are thus all  
 CC potentially involved in catalytic activity, product inhibition  
 CC and/or enzyme stability. Modified hMnSODs, or expression  
 CC vectors in which a modified hMnSOD nucleic acid is linked to a  
 CC promoter (preferably mammalian), are used to protect a cell line  
 CC from damage caused by superoxide radicals (claimed). They can also  
 CC be used to treat subjects suffering from, or at risk of,  
 CC cytotoxicity caused by superoxide radicals (claimed). As such,  
 CC they can be used as antioxidants in the treatment of a variety of  
 CC disorders, including inflammation (claimed), reperfusion injury  
 CC following ischemia (claimed), and cellular damage caused by  
 CC chemotherapeutic agents.  
 CC Note. The present sequence is not shown in the specification but  
 CC is derived from the hMnSOD sequence given on page 53-54.  
 XX  
 SO Sequence 222 AA;  
 Query Match 51.0%; Score 74; DB 20; Length 222;  
 Best Local Similarity 42.9%; Pred. No. 0.00069;  
 Matches 12; Conservative 10; Mismatches 6; Indels 0; Gaps 0;  
 Oy 1 KYSLPELDYFSATEPYISGOINEXYT 28  
 Db 25 KHSPLDLPYDYGALPHINAQIMQLHHS 52

RESULT 29  
 AAW98177  
 ID AAW98177 standard; Protein; 222 AA.  
 XX  
 AC AAW98177;  
 XX  
 DT 05-JUL-1999 (first entry)  
 XX  
 DE Human manganese superoxide dismutase H30S mutant.  
 XX  
 KM Superoxide dismutase; MnSOD; SOD; human; protein engineering;  
 KM enzyme engineering; antioxidant; antiinflammatory; cytotoxicity;  
 KM inflammation; reperfusion injury; therapy; mutant.  
 XX  
 OS Homo sapiens.  
 XX  
 XX Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 FT Peptide 1..24  
 FT /note= "signal peptide"  
 FT Protein 25..222  
 FT /note= "mature protein"  
 FT 50  
 FT /note= "His-26"  
 FT 54  
 FT /note= "His-30"  
 FT 58  
 FT /note= "Tyr-58"  
 FT 98  
 FT /note= "Tyr-74"  
 FT 167  
 FT /note= "Gln-143"  
 FT 183  
 FT /note= "Asp-159"  
 FT 185  
 FT /note= "Trp-161"  
 FT 187  
 FT /note= "His-163"  
 FT 155  
 FT /label= Gln, Glu  
 FT Misc-difference 54  
 FT /note= "His replaces Ser in wild-type hMnSOD"  
 FT  
 XX  
 XX W09913088-A1.  
 XX  
 XX 18-MAR-1999.  
 XX  
 XX PF 10-SEP-1998; 98MO-US18842.  
 XX  
 XX PR 10-SEP-1997; 97US-0927230.  
 XX  
 XX (UYFL ) UNIV FLORIDA.  
 XX  
 XX Nick HS, Silverman DN;  
 XX  
 XX WPI; 1999-229242/19.  
 XX  
 XX New recombinant human manganese superoxide dismutase proteins  
 XX  
 XX Claim 33; Page -; 61pp; English.  
 XX  
 XX This protein is a H30S mutant of human manganese superoxide  
 CC dismutase (hMnSOD, see also AAW98170). Novel modified hMnSOD  
 CC proteins having catalytic activity which differs from natural  
 CC hMnSOD are claimed. The modified proteins (see also AAW98172-77)  
 CC exhibit reduced or no product inhibition, or have greater  
 CC activity, or both, compared to natural hMnSOD. The modifications  
 CC involve one or amino acid substitutions within the active site of  
 CC the enzyme, especially at residues His-26, His-30, Tyr-34, His-74,  
 CC Gln-143, Asp-159, Trp-161 and His-163. Gln-143, Tyr-34 and His-30  
 CC are located near the active site metal, are highly conserved, and

CC are involved in a hydrogen bond relay which links the active site  
 CC metal-bound hydroxyl group to ordered solvent, and are thus all  
 CC potentially involved in catalytic activity, product inhibition  
 CC and/or enzyme stability. Modified hMnSODs, or expression  
 CC vectors in which a modified hMnSOD nucleic acid is linked to a  
 CC promoter (preferably mammalian), are used to protect a cell line  
 CC from damage caused by superoxide radicals (claimed). They can also  
 CC be used to treat subjects suffering from, or at risk of,  
 CC cytotoxicity caused by superoxide radicals (claimed). As such,  
 CC they can be used as antioxidants in the treatment of a variety of  
 CC disorders, including inflammation (claimed), reperfusion injury  
 CC following ischemia (claimed), and cellular damage caused by  
 CC chemotherapeutic agents.  
 CC Note. The present sequence is not shown in the specification but  
 CC is derived from the hMnSOD sequence given on page 53-54.  
 XX  
 SQ Sequence 222 AA;  
 XX  
 Query Match 51.0%; Score 74; DB 20; Length 222;  
 Best Local Similarity 42.3%; Pred. No. 0.00069;  
 Matches 12; Conservative 10; Mismatches 6; Indels 0; Gaps 0;  
 QY 1 KYSLPELDYERSAREPYISQGINEIXYT 28  
 DB 25 KHSLPDLPYDYGALPEPHINQIQLHRS 52  
 XX  
 RESULT 30  
 AAW98169  
 ID AAW98169 standard; Protein; 222 AA.  
 XX  
 AC AAW98169;  
 XX  
 DT 05-JUL-1999 (first entry)  
 XX  
 DE Human native manganese superoxide dismutase.  
 XX  
 KM Superoxide dismutase; MnSOD; SOD; human; protein engineering;  
 KM enzyme engineering; antioxidant; antiinflammatory; cytotoxicity;  
 KM inflammation; reperfusion injury; therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FT Peptide 1..24  
 FT /note= "signal peptide"  
 FT Protein 25..222  
 FT /note= "mature protein"  
 FT 50  
 FT /note= "His-26"  
 FT 54  
 FT /note= "His-30"  
 FT 58  
 FT /note= "Tyr-58"  
 FT 98  
 FT /note= "Tyr-74"  
 FT 167  
 FT /note= "Gln-143"  
 FT 183  
 FT /note= "Asp-159"  
 FT 185  
 FT /note= "Trp-161"  
 FT 187  
 FT /note= "His-163"  
 FT Active-site  
 XX  
 XX W09913088-A1.  
 XX  
 XX 18-MAR-1999.  
 XX  
 XX PF 10-SEP-1998; 98MO-US18842.  
 XX  
 XX PR 10-SEP-1997; 97US-0927230.  
 XX

[illegible]

FT	Active-site	/note= "Trp-161"
FT	Active-site	187
FT	Misc-difference	/note= "His-163"
FT		155
XX		/label= Gln, Glu
PN	MO9913088-A1.	
XX	18-MAR-1999.	
XX	10-SEP-1998;	98WO-US18842.
PF	10-SEP-1997;	97US-0927230.
PR	(UYFL ) UNIV FLORIDA.	
PA	Nick HS, Silverman DN;	
P1	WPI; 1999-229242/19.	
DR	New recombinant human manganese superoxide dismutase proteins	
PT	Claim 2; Page 53-54; 61gp; English.	
XX	This protein is human manganese superoxide dismutase (hmSOD).	
CC	Novel hmSOD proteins having catalytic activity which differs from	
CC	this natural hmSOD are claimed (see AAW98171-77). The modified	
CC	proteins exhibit reduced or no product inhibition, or have greater	
CC	activity, or both, compared to natural hmSOD. The modifications	
CC	involve one or amino acid substitutions within the active site of	
CC	the enzyme, especially at residues His-26, His-30, Tyr-34, His-74,	
CC	Gln-143, Asp-159, Trp-161 and His-163. The modified hmSOD, or	
CC	expression vectors in which modified hmSOD nucleic acid is linked	
CC	to a promoter (preferably mammalian), can be used to protect a cell	
CC	line from damage caused by superoxide radicals (claimed). They can	
CC	also be used to treat subjects suffering from, or at risk of,	
CC	cytotoxicity caused by superoxide radicals (claimed). As such,	
CC	they can be used as antioxidants in the treatment of a variety of	
CC	disorders, including inflammation (claimed), reperfusion injury	
CC	following ischemia (claimed), and cellular damage caused by	
CC	chemotherapeutic agents.	
XX		
SQ	Sequence 222 AA;	
	Query Match	51.0%; Score 74; DB 20; Length 222;
	Best Local Similarity	42.9%; Pred. No. 0.00069;
	Matches 12; Conservative 10; Mismatches 6; Indels 0; Gaps 0;	
QY	1 KYSLPELDYFSATEPYISGQINEIXYT 28	
	:       :   :   :   :   :   :   :	
DB	25 KHSLPDLFDYGALPHTNAIQIMQLHHS 52	
ID	AAW98171 standard; Protein; 222 AA.	
XX	AAW98171;	
AC	AAW98171;	
DT	05-JUL-1999 (first entry)	
XX	Human manganese superoxide dismutase Q143N mutant.	
DE		
KX	Superoxide dismutase; MnSOD; SOD; human; protein engineering;	
KM	enzyme engineering; antioxidant; antiinflammatory; cytotoxicity;	
KW	Inflammation; reperfusion injury; therapy; mutant.	
XX		
OS	Homo sapiens.	
OS	Synthetic.	
XX		
TH	Key	Location/Qualifiers
PT	Peptide	1..24







FT /note= "mature protein"  
 XX US6326003-B1.  
 XX  
 PD 04-DEC-2001.  
 XX  
 PF 14-OCT-1986; 86US-0918534.  
 XX  
 PR 14-OCT-1986; 86US-0918534.  
 XX  
 PA (CHIR ) CHIRON CORP.  
 XX  
 PI Halliwell RA, Bell GI, Mullenbach GT;  
 XX  
 DR WPI; 2002-129495/17.  
 DR N-PSDB; ABA94453.  
 XX  
 PT Treating inflammatory joint disease such as osteoarthritis, rheumatoid  
 PT arthritis and post ischemic tissue injury, comprises injecting human  
 PT manganese superoxide dismutase -  
 XX  
 PS Example 2; Fig 4; 14pp; English.  
 XX  
 CC The invention relates to a method of treating a patient with inflammatory  
 CC joint disease. The method involves intra-articularly injecting into the  
 CC afflicted joint a solution of human manganese superoxide dismutase  
 CC (hSODm). The treatment is used to treat an inflammatory joint disease,  
 CC particularly osteoarthritis or rheumatoid arthritis. The method can also  
 CC be used to treat post ischemic tissue injury. The present sequence  
 CC represents the hSODm polypeptide.  
 XX  
 SQ Sequence 222 AA;  
 XX  
 Query Match 51.0%; Score 74; DB 23; Length 222;  
 Best Local Similarity 42.9%; Pred. No. 0.00069;  
 Matches 12; Conservative 10; Mismatches 6; Indels 0; Gaps 0;  
 QY 1 KYSLPELDYEFSGATEPYISQGINEIXYT 28  
 Db 25 KHSLEPDLPEYDYGALPEPHINQIMQLHHS 52  
 XX  
 RESULT 39  
 AAP80551  
 ID AAP80551 standard; protein; 223 AA.  
 XX  
 AC AAP80551;  
 XX  
 DT 22-OCT-1990 (first entry)  
 XX  
 DE Human manganese superoxide dismutase.  
 XX  
 KM Human manganese superoxide dismutase; oxygen free radicals.  
 KM  
 XX Homo sapiens.  
 OS  
 XX EP284105-A.  
 PN  
 XX 28-SEP-1988.  
 PD  
 XX 25-MAR-1988; 88EP-0104880.  
 PF  
 XX 26-FEB-1988; 88US-0161117.  
 PR  
 PR 27-MAR-1987; 87US-0032734.  
 XX  
 PA (BIOT-) BIO-TECHNOLOGY GENERAL CORPORATION.  
 XX  
 PI Hartman JR, Beck Y, Nimrod A;  
 XX  
 DR WPI; 1988-272584/39.  
 DR N-PSDB; AAN81158.  
 XX  
 PT Recombinant human manganese superoxidisedismutase -

PT used for treating, e.g. reperfusion injury, inflammation,  
 PT arthritis, bronchial pulmonary dysplasia or lung fibrosis.  
 XX  
 PS Disclosure; ; pp; English.  
 XX  
 CC Human manganese superoxide dismutase or its analogues or mutants is  
 CC used to treat subjects suffering from a disorder associated with the  
 CC generation of oxygen free radicals, eg reperfusion injury following  
 CC ischaemia or organ transplant, inflammation caused by Crohn's disease  
 CC or colitis, synovial inflammation induced by bacterial lipopolysaccharide  
 CC endotoxin, arthritis, bronchial pulmonary dysplasia or lung fibrosis  
 CC by radiation or an anti-neoplastic agent, eg bleomycin. It can also be  
 CC used to prolong the life of excised organs. As it remains in the serum  
 CC for long periods, it may be used as a long-acting therapeutic agent.  
 CC Dosage is 1-200 mg/joint or 1-100, esp. 3-50, mg/kg body wt. daily.  
 CC It is given intra-articularly, intratracheally, sc or iv.  
 XX  
 SQ Sequence 223 AA;  
 XX  
 Query Match 51.0%; Score 74; DB 9; Length 223;  
 Best Local Similarity 42.9%; Pred. No. 0.0007;  
 Matches 12; Conservative 10; Mismatches 6; Indels 0; Gaps 0;  
 QY 1 KYSLPELDYEFSGATEPYISQGINEIXYT 28  
 Db 26 KHSLEPDLPEYDYGALPEPHINQIMQLHHS 53  
 XX  
 RESULT 40  
 AAY55849  
 ID AAY55849 standard; protein; 226 AA.  
 XX  
 AC AAY55849;  
 XX  
 DT 06-MAR-2000 (first entry)  
 XX  
 DE Human manganese superoxide dismutase (MnSOD) mutant protein.  
 XX  
 KM Manganese superoxide dismutase; MnSOD; ECSOD; proteoglycan; human;  
 KM endothelial; oxidative damage; pulmonary inflammatory injury; cancer;  
 KM lung disease; hypoxia; ischemia; reperfusion injury; arthritis;  
 KM hyperoxia; atherosclerosis; lupus erythematosus; hypertension;  
 KM neutrophil-mediated inflammation; mutant.  
 XX  
 OS Synthetic.  
 OS  
 XX Homo sapiens.  
 OS  
 XX WO958547-A1.  
 PN  
 XX 18-NOV-1999.  
 PD  
 XX 06-MAY-1999; 99WO-US09921.  
 PF  
 XX 08-MAY-1998; 98US-0075019.  
 PR  
 XX (WEBB-) WEBB-WARING INST BIOMEDICAL RES.  
 XX  
 PI McCord JM, Gao B, Flores SC;  
 XX  
 XX WPI; 2000-062283/05.  
 DR  
 DR N-PSDB; AAZ39781, AAZ39782.  
 XX  
 PT Modified manganese superoxide dismutase, methods of production and  
 PT antibodies -  
 XX  
 PS Claim 29; Page 74-75; 83pp; English.  
 XX  
 CC The invention provides a nucleic acid molecule encoding a genetically  
 CC modified manganese superoxide dismutase (MnSOD). The nucleic acid  
 CC comprises: a first nucleic acid sequence encoding an enzymatically active  
 CC portion of MnSOD; and a second nucleic acid sequence encoding a peptide  
 CC (ECSOD) that binds to polyanionic polysaccharides or proteoglycans on  
 CC endothelial cell surfaces. The protein protects a mammal, especially







GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 9, 2003, 14:00:55 ; Search time 14.3119 Seconds

(without alignments)  
279.314 Million cell updates/sec

Title: US-09-987-190-2

Perfect score: 30

Sequence: 1 KYSLPELDYEFSAATEPYISQGINEIXYTX 30

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 908470 seqs, 133250620 residues

Word size : 0

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database : A\_Geneseq\_101002.\*

1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	90.0	30	19	AAAG3252
2	10	33.3	23	23	ABBO5612
3	7	23.3	407	21	AAAG14176
4	7	23.3	512	21	AAAG14175
5	7	23.3	521	21	AAAG14174
6	6	20.0	7	23	ABBO8465
7	6	20.0	7	23	ABBO8476
8	6	20.0	9	23	ABBO8466
9	6	20.0	10	23	ABBO8467
10	6	20.0	19	23	ABG66263

11	6	20.0	20	23	ABBO8468	Mouse tumour speci
12	6	20.0	44	22	AAAB7878	C. glutamicum SKT
13	6	20.0	50	22	ABAB36416	Peptide #3922 enco
14	6	20.0	50	22	AAAB57176	Human brain expres
15	6	20.0	50	22	AAAB69579	Human bone marrow
16	6	20.0	50	22	AAAB73397	Peptide #3831 enco
17	6	20.0	50	22	AAAB29917	Human peptide enco
18	6	20.0	52	22	ABG39201	Propionibacterium
19	6	20.0	52	22	AAUB58434	Propionibacterium
20	6	20.0	52	22	AAUB5746	Human nervous syst
21	6	20.0	62	23	ABAB4809	Human tumour speci
22	6	20.0	62	23	ABBO8458	Human secreted pro
23	6	20.0	81	19	AAAB75089	Human polypeptide
24	6	20.0	136	22	AAO01578	Human prostate can
25	6	20.0	146	21	AAAB56397	Streptococcus poly
26	6	20.0	163	23	ABBP26232	Candida albicans f
27	6	20.0	188	19	AAAB53253	C. glutamicum prote
28	6	20.0	200	22	AAAB32973	C. melanoscola sup
29	6	20.0	200	22	AAAB47790	Listeria monocytog
30	6	20.0	207	23	AAAB47790	C. pneumoniae prot
31	6	20.0	214	20	AAAB4657	Human protein sequ
32	6	20.0	221	22	AAAB94056	Piscicirickettsia sa
33	6	20.0	226	22	AAAB78026	Arabidopsis thalia
34	6	20.0	294	21	AAAB49754	Arabidopsis thalia
35	6	20.0	294	21	AAAB50274	Arabidopsis thalia
36	6	20.0	311	21	AAAB29351	Arabidopsis thalia
37	6	20.0	312	21	AAAB49753	Arabidopsis thalia
38	6	20.0	312	21	AAAB50273	Arabidopsis thalia
39	6	20.0	338	23	ABAB4867	Lactococcus lactis
40	6	20.0	355	21	AAAB29350	Arabidopsis thalia
41	6	20.0	384	21	AAAB29350	Arabidopsis thalia
42	6	20.0	444	22	AAAB49921	Shrimp white spot
43	6	20.0	454	21	AAAB53794	Arabidopsis thalia
44	6	20.0	487	21	AAAB49752	Arabidopsis thalia
45	6	20.0	506	21	AAAB29349	Arabidopsis thalia
46	6	20.0	507	21	AAAB50272	Arabidopsis thalia
47	6	20.0	511	21	AAAB53793	Arabidopsis thalia
48	6	20.0	541	19	AAAB44165	Entodinium caudatu
49	6	20.0	545	21	AAAB53792	Arabidopsis thalia
50	6	20.0	620	22	ABAB6065	Drosophila melanog
51	6	20.0	640	23	ABAB3563	Lactococcus lactis
52	6	20.0	650	22	ABAB1781	Drosophila melanog
53	6	20.0	729	22	ABAB58260	Drosophila melanog
54	6	20.0	901	22	AAAB97741	H. annus LOX prote
55	6	20.0	934	23	ABAB34123	Human protein from
56	6	20.0	1023	22	AAAB11436	D. discoideum ster
57	6	20.0	1596	23	AAAB20271	Human lung, specifi
58	6	16.7	5	23	ABAB66201	Phage display g3 l
59	5	16.7	9	17	AAAB49494	Human leucocyte an
60	6	16.7	9	22	AAAB23850	Human MHC class I
61	6	16.7	9	22	AAAB24148	Human MHC molecule
62	6	16.7	9	22	AAAB24221	Human MHC class I
63	6	16.7	9	22	AAAB21507	Human HIV-1 Th-CTL
64	6	16.7	9	22	AAAB49390	Simian immunodef
65	5	16.7	9	23	ABAB34938	Human bone marrow
66	6	16.7	10	22	AAAB23867	Human MHC class I
67	6	16.7	10	22	AAAB23986	Human MHC class I
68	6	16.7	10	22	AAAB24069	Human MHC class I
69	6	16.7	10	22	AAAB24179	Human MHC molecule
70	6	16.7	11	12	AAAB10552	Residues 429-444 o
71	6	16.7	12	17	AAAB07947	gpi20 peptide 1029
72	6	16.7	13	21	AAAB58866	hMSH mutant pepid
73	6	16.7	14	22	AAAB98507	Human peptide #178
74	6	16.7	14	22	AAAB98508	Human peptide #178
75	5	16.7	15	19	AAAB76955	Fusion immunoglob
76	5	16.7	15	19	AAAB76952	Fusion immunoglob
77	5	16.7	15	23	AAAB10681	Peptide A155 conta
78	5	16.7	16	17	AAAB12640	SH2 binding peptid
79	5	16.7	17	10	AAAB91153	Artificial peptide
80	5	16.7	18	21	AAAB79501	Eimeria tenella SO
81	5	16.7	18	23	ABAB66234	IGE fcepsion RI b
82	5	16.7	19	17	AAAB07951	gpi20 peptide DW-3
83	5	16.7	19	23	ABAB66236	IGE fcepsion RI b

84	5	16.7	19	23	AAU75606	Human type IV coll	157	5	16.7	21	23	ABG66625	Phage displayed g3
85	5	16.7	20	18	AAW42955	Immunogenic Hepat	158	5	16.7	21	23	ABG66626	Phage displayed g3
86	5	16.7	20	23	ABG66233	IGF Fccepillon Ri b	159	5	16.7	21	23	ABG66627	Phage displayed g3
87	5	16.7	20	23	ABG66642	Phage displayed g3	160	5	16.7	21	23	ABG66629	Phage displayed g3
88	5	16.7	21	22	ABG69475	Synthetic HAV p3d	161	5	16.7	21	23	ABG66637	Phage displayed g3
89	5	16.7	21	23	ABG66200	Phage displayed g3	162	5	16.7	21	23	ABG66638	Phage displayed g3
90	5	16.7	21	23	ABG66202	IGF Fccepillon Ri b	163	5	16.7	21	23	ABG66639	Phage displayed g3
91	5	16.7	21	23	ABG66218	IGF Fccepillon Ri b	164	5	16.7	21	23	ABG66640	Phage displayed g3
92	5	16.7	21	23	ABG66221	IGF Fccepillon Ri b	165	5	16.7	21	23	ABG66641	Phage displayed g3
93	5	16.7	21	23	ABG66252	IGF Fccepillon Ri b	166	5	16.7	23	21	AAH11664	Peptide 1017/88 ba
94	5	16.7	21	23	ABG66253	IGF Fccepillon Ri b	167	5	16.7	23	21	AAV70282	Plasmodium falcipa
95	5	16.7	21	23	ABG66254	IGF Fccepillon Ri b	168	5	16.7	23	23	ABG66625	IGF Fccepillon Ri b
96	5	16.7	21	23	ABG66264	IGF Fccepillon Ri b	169	5	16.7	23	23	ABG66626	IGF Fccepillon Ri b
97	5	16.7	21	23	ABG66273	IGF Fccepillon Ri b	170	5	16.7	24	12	AAH11630	Peptide 1029/16 ba
98	5	16.7	21	23	ABG66359	IGF Fccepillon Ri b	171	5	16.7	24	14	AAH33837	Cluster peptide PC
99	5	16.7	21	23	ABG66460	IGF Fccepillon Ri b	172	5	16.7	24	15	AAH33845	HIV-1 gp120-35 epi
100	5	16.7	21	23	ABG66461	IGF Fccepillon Ri b	173	5	16.7	24	16	AAH66431	PCTUS 3-18 (421-44
101	5	16.7	21	23	ABG66462	IGF Fccepillon Ri b	174	5	16.7	24	20	AAV05353	HIV-1 CUVAC pep1
102	5	16.7	21	23	ABG66463	IGF Fccepillon Ri b	175	5	16.7	24	23	ABG66627	IGF Fccepillon Ri b
103	5	16.7	21	23	ABG66464	IGF Fccepillon Ri b	176	5	16.7	24	23	ABG66658	IGF Fccepillon Ri b
104	5	16.7	21	23	ABG66465	IGF Fccepillon Ri b	177	5	16.7	24	23	ABG66659	IGF Fccepillon Ri b
105	5	16.7	21	23	ABG66480	IGF Fccepillon Ri b	178	5	16.7	24	23	ABG66591	IGF Fccepillon Ri b
106	5	16.7	21	23	ABG66510	IGF Fccepillon Ri b	179	5	16.7	24	23	ABG66592	IGF Fccepillon Ri b
107	5	16.7	21	23	ABG66511	IGF Fccepillon Ri b	180	5	16.7	24	23	ABG66593	IGF Fccepillon Ri b
108	5	16.7	21	23	ABG66512	IGF Fccepillon Ri b	181	5	16.7	24	23	ABG66594	IGF Fccepillon Ri b
109	5	16.7	21	23	ABG66513	IGF Fccepillon Ri b	182	5	16.7	24	23	ABG66595	IGF Fccepillon Ri b
110	5	16.7	21	23	ABG66514	IGF Fccepillon Ri b	183	5	16.7	24	23	ABG66596	IGF Fccepillon Ri b
111	5	16.7	21	23	ABG66515	IGF Fccepillon Ri b	184	5	16.7	24	23	ABG66597	IGF Fccepillon Ri b
112	5	16.7	21	23	ABG66516	IGF Fccepillon Ri b	185	5	16.7	25	8	AAH71144	Anti-HIV peptide 5
113	5	16.7	21	23	ABG66517	IGF Fccepillon Ri b	186	5	16.7	25	9	AAH80749	Sequence of peptid
114	5	16.7	21	23	ABG66518	IGF Fccepillon Ri b	187	5	16.7	25	13	AAH27331	Peptide corresp. c
115	5	16.7	21	23	ABG66519	IGF Fccepillon Ri b	188	5	16.7	26	20	AAH25833	Human secreted pro
116	5	16.7	21	23	ABG66520	IGF Fccepillon Ri b	189	5	16.7	26	23	ABG66272	IGF Fccepillon Ri b
117	5	16.7	21	23	ABG66521	IGF Fccepillon Ri b	190	5	16.7	26	23	ABG66598	IGF Fccepillon Ri b
118	5	16.7	21	23	ABG66522	IGF Fccepillon Ri b	191	5	16.7	26	23	ABG66599	IGF Fccepillon Ri b
119	5	16.7	21	23	ABG66523	IGF Fccepillon Ri b	192	5	16.7	26	23	ABG66600	IGF Fccepillon Ri b
120	5	16.7	21	23	ABG66524	IGF Fccepillon Ri b	193	5	16.7	26	23	ABG66601	IGF Fccepillon Ri b
121	5	16.7	21	23	ABG66525	IGF Fccepillon Ri b	194	5	16.7	26	23	ABG66602	IGF Fccepillon Ri b
122	5	16.7	21	23	ABG66526	IGF Fccepillon Ri b	195	5	16.7	27	15	AAH63844	HIV-1 gp120-34 epi
123	5	16.7	21	23	ABG66527	IGF Fccepillon Ri b	196	5	16.7	28	11	AAH03100	Human immunodefict
124	5	16.7	21	23	ABG66528	IGF Fccepillon Ri b	197	5	16.7	28	11	AAH09346	Sequence of centra
125	5	16.7	21	23	ABG66529	IGF Fccepillon Ri b	198	5	16.7	28	12	AAH11631	Peptide 1005/45 ba
126	5	16.7	21	23	ABG66530	IGF Fccepillon Ri b	199	5	16.7	28	13	AAH23032	Cyclic 1005-45 pep
127	5	16.7	21	23	ABG66531	IGF Fccepillon Ri b	200	5	16.7	28	16	AAH67718	HIV isolate LAI-II
128	5	16.7	21	23	ABG66564	IGF Fccepillon Ri b	201	5	16.7	28	19	AAH54910	HIV gp120 envelope
129	5	16.7	21	23	ABG66565	IGF Fccepillon Ri b	202	5	16.7	28	19	AAH54911	HIV gp120 envelope
130	5	16.7	21	23	ABG66566	IGF Fccepillon Ri b	203	5	16.7	28	19	AAH54914	HIV gp120 envelope
131	5	16.7	21	23	ABG66567	IGF Fccepillon Ri b	204	5	16.7	28	19	AAH54915	HIV gp120 envelope
132	5	16.7	21	23	ABG66568	IGF Fccepillon Ri b	205	5	16.7	28	19	AAH54916	HIV gp120 envelope
133	5	16.7	21	23	ABG66569	IGF Fccepillon Ri b	206	5	16.7	29	13	AAH27330	Peptide corresp. c
134	5	16.7	21	23	ABG66570	IGF Fccepillon Ri b	207	5	16.7	29	15	AAH58550	HIV env CD4 bindin
135	5	16.7	21	23	ABG66571	IGF Fccepillon Ri b	208	5	16.7	29	22	AAH12470	HIV Th-CTL peptide
136	5	16.7	21	23	ABG66572	IGF Fccepillon Ri b	209	5	16.7	29	22	AAH49394	Simian immunodefic
137	5	16.7	21	23	ABG66573	IGF Fccepillon Ri b	210	5	16.7	29	23	AAH70233	HIV/SIV TH-CTL epi
138	5	16.7	21	23	ABG66574	IGF Fccepillon Ri b	211	5	16.7	32	19	AAH76910	Fusion immunoglobu
139	5	16.7	21	23	ABG66575	IGF Fccepillon Ri b	212	5	16.7	33	23	AAH48734	Human peptide enco
140	5	16.7	21	23	ABG66576	IGF Fccepillon Ri b	213	5	16.7	34	22	AAH84873	Mousses liver growth
141	5	16.7	21	23	ABG66577	IGF Fccepillon Ri b	214	5	16.7	34	23	AAH10677	Peptide T35S conta
142	5	16.7	21	23	ABG66578	IGF Fccepillon Ri b	215	5	16.7	35	12	AAH13121	Binding site of G3
143	5	16.7	21	23	ABG66579	IGF Fccepillon Ri b	216	5	16.7	36	12	AAH14592	HIV gp120-derived
144	5	16.7	21	23	ABG66580	IGF Fccepillon Ri b	217	5	16.7	36	12	AAH64773	Human secreted pro
145	5	16.7	21	23	ABG66581	IGF Fccepillon Ri b	218	5	16.7	37	22	AAH63980	Amino acid sequenc
146	5	16.7	21	23	ABG66582	IGF Fccepillon Ri b	219	5	16.7	39	16	AAH66445	PCTUS 3-18 (421-44
147	5	16.7	21	23	ABG66583	IGF Fccepillon Ri b	220	5	16.7	39	16	AAH66446	PCTUS 3-18 (421-44
148	5	16.7	21	23	ABG66584	IGF Fccepillon Ri b	221	5	16.7	39	19	AAH66088	S. pneumoniae deri
149	5	16.7	21	23	ABG66585	IGF Fccepillon Ri b	222	5	16.7	39	20	AAH05346	HIV-1 CUVAC pep1
150	5	16.7	21	23	ABG66586	IGF Fccepillon Ri b	223	5	16.7	39	20	AAH05339	HIV-1 CUVAC pep1
151	5	16.7	21	23	ABG66587	IGF Fccepillon Ri b	224	5	16.7	39	22	AAH28435	Peptide #1086 enco
152	5	16.7	21	23	ABG66617	Phage displayed g3	225	5	16.7	39	22	AAH33617	Peptide #1123 enco
153	5	16.7	21	23	ABG66619	Phage displayed g3	226	5	16.7	39	22	AAH19072	Protein #1071 enco
154	5	16.7	21	23	ABG66620	Phage displayed g3	227	5	16.7	39	22	AAH07857	Synthetic peptide
155	5	16.7	21	23	ABG66621	Phage displayed g3	228	5	16.7	39	22	AAH54350	Human brain expres
156	5	16.7	21	23	ABG66624	Phage displayed g3	229	5	16.7	39	22	AAH66791	Human Bone marrow

230	5	16.7	39	22	AAAI4657	Peptide #1091 enco	303	5	16.7	83	22	AAU23141	Novel human enzyme
231	5	16.7	39	22	AAAI2706	Peptide #1113 enco	304	5	16.7	84	21	AAAG15357	Arabidopsis thalia
232	5	16.7	39	22	AAAM02380	Peptide #1062 enco	305	5	16.7	84	21	ABG26358	Novel human diagno
233	5	16.7	39	23	ABG68641	HIV-1 gp160 V3 loc	306	5	16.7	85	22	AAU14747	Novel bone marrow
234	5	16.7	39	23	ABG68648	HIV-1 gp160 V3 loc	307	5	16.7	86	23	AAU10479	HIV/Siv TH-CTL epi
235	5	16.7	39	23	ABG36446	Human peptide enco	308	5	16.7	87	21	AAAG36837	Arabidopsis thalia
236	5	16.7	40	20	AAW93231	Human cytochrome p	309	5	16.7	88	22	AAAG2974	Human digestive sy
237	5	16.7	40	20	AAW93232	Human cytochrome p	310	5	16.7	88	22	AAU20070	Human liver associ
238	5	16.7	41	12	ABG22627	Novel human diagno	311	5	16.7	88	23	ABP40931	Human liver antige
239	5	16.7	44	14	AAAR41470	HIV-1 gp160/120 fr	312	5	16.7	88	23	AAU75607	Human type IV coll
240	5	16.7	45	15	AAAR60792	HIV virus-1 gp120	313	5	16.7	88	23	AAU75608	Human type IV coll
241	5	16.7	46	22	ABBS95935	Human testicular a	314	5	16.7	90	5	AAAP81887	Sequence of human
242	5	16.7	46	22	AAW95238	Human reproductive	315	5	16.7	90	21	AAAG36836	Arabidopsis thalia
243	5	16.7	49	7	AAAP61106	Protein encoded by	316	5	16.7	90	21	AAAG02990	Human secreted pro
244	5	16.7	49	22	AAE04847	Human SGP014 phosp	317	5	16.7	91	22	ABBA40418	Peptide #7924 enco
245	5	16.7	51	22	ABG29596	Novel human diagno	318	5	16.7	91	22	ABBA24781	Protein #6780 enco
246	5	16.7	52	21	AAAG58752	Arabidopsis thalia	319	5	16.7	91	22	AAAG61225	Human brain expres
247	5	16.7	52	22	AAU18951	Novel lung cancer	320	5	16.7	91	22	AAAG73945	Human bone marrow
248	5	16.7	52	22	AAU17963	Novel human respit	321	5	16.7	91	22	AAAG41423	Human type IV coll
249	5	16.7	54	16	AAAR71644	Circumsporozoit R	322	5	16.7	92	22	ABBI0771	Human peptide enco
250	5	16.7	54	16	AAU14653	Novel bone marrow	323	5	16.7	92	22	ABG43833	Human ovarian and/
251	5	16.7	55	23	ABP07553	Human ORFX protein	324	5	16.7	92	22	AAAG94434	Human reproductive
252	5	16.7	56	7	AAAP61109	Protein encoded by	325	5	16.7	92	22	AAAG96608	Human immune/haema
253	5	16.7	56	7	AAAP61107	Protein encoded by	326	5	16.7	93	22	AAO00268	Human polypeptide
254	5	16.7	56	23	ABP33744	Human ORF2717 prot	327	5	16.7	94	20	AAAY29670	Human src-family k
255	5	16.7	57	22	AAU039400	Protonibacterium	328	5	16.7	94	22	AAU08732	Src-family kinase
256	5	16.7	57	23	ABP08560	Human ORFX protein	329	5	16.7	94	22	AAAG21447	Human polypeptide
257	5	16.7	58	12	AAAR14591	HIV gp120-derived	330	5	16.7	94	23	ABP34090	Human ORF3063 prot
258	5	16.7	58	19	AAAM40522	Chicken p75 neurot	331	5	16.7	95	22	AAAG89809	Human immune/haema
259	5	16.7	58	19	AAAM40520	Human p75 neurot	332	5	16.7	95	23	ABP02403	Human ORFX protein
260	5	16.7	58	19	AAAM40521	Rat p75 neurotrophin	333	5	16.7	96	22	AAU62138	Protonibacterium
261	5	16.7	60	22	ABBI14861	Human nervous syst	334	5	16.7	97	21	AAAB25332	Eucalyptus grandis
262	5	16.7	61	20	AAAY39194	M. tuberculosis an	335	5	16.7	99	21	AAAG57852	Arabidopsis thalia
263	5	16.7	61	20	AAAY39051	M. tuberculosis re	336	5	16.7	99	21	AAAG59853	Arabidopsis thalia
264	5	16.7	61	23	ABP34409	Human ORF3382 prot	337	5	16.7	99	21	AAAG60252	Arabidopsis thalia
265	5	16.7	62	21	AAAG15941	Arabidopsis thalia	338	5	16.7	99	21	AAU41923	Protonibacterium
266	5	16.7	62	22	AAU40469	Protonibacterium	339	5	16.7	99	23	ABP28999	Streptococcus poly
267	5	16.7	63	22	AAAM5133	Human novel foetal	340	5	16.7	100	23	ABBB89123	Human polypeptide
268	5	16.7	63	22	AAAM5133	Human immune/haema	341	5	16.7	101	23	ABBP6470	Human ORFX protein
269	5	16.7	64	22	AAAG5535	Human reproductive	342	5	16.7	102	21	AAAG04069	Human secreted pro
270	5	16.7	65	22	ABBB28619	Peptide #1270 enco	343	5	16.7	102	21	AAAG04086	Human secreted pro
271	5	16.7	65	22	ABBB3802	Peptide #1308 enco	344	5	16.7	102	23	ABP34257	Human structural p
272	5	16.7	65	22	ABBI19245	Protein #1144 enco	345	5	16.7	102	23	ABBP04321	Human ORFX protein
273	5	16.7	65	22	AAAM4568	Human brain expres	346	5	16.7	104	23	ABBP6168	Human ORFX protein
274	5	16.7	65	22	AAAM66975	Human bone marrow	347	5	16.7	105	22	ABBB03916	Human musculoskele
275	5	16.7	65	22	AAAM4835	Peptide #1269 enco	348	5	16.7	105	23	ABBB02927	Human polypeptide
276	5	16.7	65	22	AAAM7262	Peptide #1299 enco	349	5	16.7	106	22	AAU41827	Protonibacterium
277	5	16.7	65	22	AAAM02560	Peptide #1242 enco	350	5	16.7	108	21	AAAY64613	Nonclassical cadhe
278	5	16.7	66	23	ABG36632	Human peptide enco	351	5	16.7	109	23	ABBP05073	Human Abeta bindin
279	5	16.7	66	16	AAAR66442	Plasmodium falcipa	352	5	16.7	111	21	AAAY97159	Micromospora eve
280	5	16.7	66	20	AAAY00258	Human secreted pro	353	5	16.7	111	22	AAU04889	Human alpha2 macro
281	5	16.7	66	22	AAU44897	Protonibacterium	354	5	16.7	111	23	AAU81024	Human alpha2 macro
282	5	16.7	66	22	AAAM1076	Human immune/haema	355	5	16.7	115	13	AAAR29911	Type C natriuretic
283	5	16.7	68	15	AAAR47845	HIV gp120 coat pro	356	5	16.7	116	19	AAAM80401	A secreted protein
284	5	16.7	69	22	AAAG75957	Human colon cancer	357	5	16.7	116	21	AAAB32655	Eucalyptus grandis
285	5	16.7	70	22	ABG27298	Novel human diagno	358	5	16.7	117	23	ABBP32154	Human ORF1127 prot
286	5	16.7	71	22	ABG12595	Novel human diagno	359	5	16.7	120	18	AAU67985	Staphylococcus aur
287	5	16.7	73	23	ABP11128	Human ORFX protein	360	5	16.7	120	22	AAU63953	Protonibacterium
288	5	16.7	74	23	ABP32966	Human ORF1399 prot	361	5	16.7	120	22	AAU23179	Novel human enzyme
289	5	16.7	75	22	AAAM83306	Human immune/haema	362	5	16.7	120	22	AAU21843	Novel human neopla
290	5	16.7	76	20	AAAM94643	TNF-R extracellular	363	5	16.7	121	21	AAAG36835	Arabidopsis thalia
291	5	16.7	76	22	AAAB69195	Human TNF-R extrac	364	5	16.7	123	21	AAAG27109	Zea mays protein f
292	5	16.7	77	22	ABBA1024	Peptide #8530 enco	365	5	16.7	123	23	AAU75554	Human type IV coll
293	5	16.7	78	20	AAAY59900	Human normal pancr	366	5	16.7	125	22	AAU27552	Human G-Protein co
294	5	16.7	78	20	AAAY59919	Human normal pancr	367	5	16.7	126	21	AAAG57751	Arabidopsis thalia
295	5	16.7	78	22	AAU64977	Protonibacterium	368	5	16.7	126	21	AAAG60251	Arabidopsis thalia
296	5	16.7	80	22	AAU75600	Human type IV coll	369	5	16.7	126	23	AAU81023	Human alpha2 macro
297	5	16.7	80	22	AAU75528	Protonibacterium	370	5	16.7	127	20	AAAY13140	Human secreted pro
298	5	16.7	80	22	ABG31296	Novel human diagno	371	5	16.7	127	22	AAU50235	Catalpa lysine dec
299	5	16.7	81	22	AAAM89740	Human immune/haema	372	5	16.7	127	23	AAU76447	AbMy-W1 transcript
300	5	16.7	81	23	AAU81025	Human alpha2 macro	373	5	16.7	129	21	AAAY58374	AbMy-W1 transcript
301	5	16.7	83	11	AAAR03106	CD4 receptor prote	374	5	16.7	129	21	AAAY58375	BDMV transcrip
302	5	16.7	83	11	AAAR09349	Sequence corresp.	375	5	16.7	129	21	AAAY58376	BDMV transcrip

376	5	16.7	129	21	AAVS6834	TMV transcription	449	5	16.7	173	22	AAAB64410	Amino acid sequenc
377	5	16.7	132	23	AAU75597	Human type IV coll	450	5	16.7	174	21	AAAG47393	Arabinidopsis thalia
378	5	16.7	133	21	AAAB25166	Eucalyptus grandis	451	5	16.7	176	13	AAAR22564	Antibody specific
379	5	16.7	133	21	AAAG19951	Arabinidopsis thalia	452	5	16.7	178	21	AAAG16324	Arabinidopsis thalia
380	5	16.7	134	21	AAAG47394	Arabinidopsis thalia	453	5	16.7	178	21	AAAG44500	Arabinidopsis thalia
381	5	16.7	134	22	ABG28264	Novel human diagno	454	5	16.7	178	21	AAAG53228	Arabinidopsis thalia
382	5	16.7	135	20	AAVS9796	Human normal ovari	455	5	16.7	179	10	AAAP92014	Arabinidopsis thalia
383	5	16.7	135	22	ABG07756	Novel human diagno	456	5	16.7	179	10	AAAP93537	HIV portion of HTL
384	5	16.7	135	22	ABBA48396	Lacteria monocyog	457	5	16.7	179	23	ABBB1829	Hericidially activ
385	5	16.7	137	18	AAW23387	Goat cytotoxic 14	458	5	16.7	180	11	AAAR07290	Circumsporozolite a
386	5	16.7	137	19	AAW71473	Amino acid sequenc	459	5	16.7	180	22	AAAG78506	Recombinant HIV en
387	5	16.7	137	19	AAW50927	Protein p14 useful	460	5	16.7	181	15	AAAR9610	HIV-1 gp120 conser
388	5	16.7	137	19	AAW56483	14 kDa protein com	461	5	16.7	181	21	AAAG06198	Arabinidopsis thalia
389	5	16.7	137	20	AAV32859	Human heat shock p	462	5	16.7	181	21	AAAG47704	Arabinidopsis thalia
390	5	16.7	137	21	AAAB36233	Goat recombinant U	463	5	16.7	181	22	ABG16191	Novel human diagno
391	5	16.7	138	22	AAAG1625	Arabinidopsis thalia	464	5	16.7	181	22	ABG29332	Novel human diagno
392	5	16.7	138	22	AAU30541	Novel human secret	465	5	16.7	182	21	ABG44356	Arabinidopsis thalia
393	5	16.7	138	22	AAAB92929	Human protein sequ	466	5	16.7	182	22	ABG16763	Novel human diagno
394	5	16.7	139	21	AAAB36234	Goat natural UK14	467	5	16.7	182	22	ABAB3656	Human protein sequ
395	5	16.7	140	21	AAAG14313	Arabinidopsis thalia	468	5	16.7	183	22	ABBS8960	Drosophila melanog
396	5	16.7	140	21	AAAG27696	Arabinidopsis thalia	469	5	16.7	185	5	AAAP40154	Sequence of human
397	5	16.7	140	21	AAAG28604	Arabinidopsis thalia	470	5	16.7	185	5	AAAP40155	Sequence of human
398	5	16.7	140	21	AAAG53204	Arabinidopsis thalia	471	5	16.7	185	10	AAAP94632	Amino acid sequenc
399	5	16.7	140	21	AAAG53209	Arabinidopsis thalia	472	5	16.7	185	23	ABG61819	Prostate cancer-as
400	5	16.7	140	21	AAAG53232	Arabinidopsis thalia	473	5	16.7	186	22	ABG15763	Novel human diagno
401	5	16.7	140	22	AAU54719	Propionibacterium	474	5	16.7	187	22	AAU87327	Novel central nerv
402	5	16.7	140	22	AAAG89765	C glutamincum prote	475	5	16.7	187	19	AAAN29746	TNF related endoth
403	5	16.7	141	21	AAV82474	Mouse ABG12 protei	476	5	16.7	189	21	AAAG46008	Arabinidopsis thalia
404	5	16.7	141	22	ABG25718	Novel human diagno	477	5	16.7	189	22	ABBS9236	Human testicular a
405	5	16.7	142	21	AAAG44357	Arabinidopsis thalia	478	5	16.7	189	22	AAU87615	Novel central nerv
406	5	16.7	142	22	ABBI12182	Human neutral prot	479	5	16.7	189	22	AAAM95239	Human reproductive
407	5	16.7	146	21	AAAG06199	Arabinidopsis thalia	480	5	16.7	189	22	AAAM3564	Human polypeptide
408	5	16.7	146	21	AAAG47705	Arabinidopsis thalia	481	5	16.7	189	22	AAAM43639	Human polypeptide
409	5	16.7	146	22	AAE00895	Human TREPA (TNP r	482	5	16.7	189	22	AAE00892	Human URAflrag TRBP
410	5	16.7	147	22	AAAB96299	Ptactive helix-tur	483	5	16.7	189	22	AAU19951	Novel human calcit
411	5	16.7	147	23	ABP41536	Human ovarian ancl	484	5	16.7	191	19	AAAG4139	Human interferon r
412	5	16.7	148	21	AAAG51579	Arabinidopsis thalia	485	5	16.7	191	19	AAAG62287	Human protein phos
413	5	16.7	148	22	AAU39237	Propionibacterium	486	5	16.7	191	19	AAAM51215	Amino acid sequenc
414	5	16.7	149	21	AAAG24388	Arabinidopsis thalia	487	5	16.7	191	21	AAAG14312	Arabinidopsis thalia
415	5	16.7	149	21	AAAG51576	Arabinidopsis thalia	488	5	16.7	191	21	AAAG7695	Arabinidopsis thalia
416	5	16.7	149	21	AAAG51582	Arabinidopsis thalia	489	5	16.7	191	21	AAAG32321	Arabinidopsis thalia
417	5	16.7	149	23	ABP28394	Streptococcus poly	490	5	16.7	191	23	AAU75596	Human type IV coll
418	5	16.7	149	23	ABBA48303	Lacteria monocyog	491	5	16.7	193	21	AAAG28603	Arabinidopsis thalia
419	5	16.7	150	15	AAAS58552	MEAV vaccine. Syn	492	5	16.7	193	21	AAAG53208	Arabinidopsis thalia
420	5	16.7	151	22	ABG21145	Novel human diagno	493	5	16.7	193	22	ABBE2640	Drosophila melanog
421	5	16.7	153	23	AAU81020	Human alpha2 macro	494	5	16.7	194	14	AAAR39483	Human apoAIV mutel
422	5	16.7	157	22	AAU25566	Human G Protein-Co	495	5	16.7	195	22	ABG00598	Novel human diagno
423	5	16.7	159	21	AAAB25331	Eucalyptus grandis	496	5	16.7	196	22	AAAG1796	C glutamicum prote
424	5	16.7	160	22	AAAM51676	Human ATP-dependen	497	5	16.7	196	22	AAAB74038	Staphylococcus aur
425	5	16.7	160	23	AAO16936	Human ATP-dependen	498	5	16.7	196	22	AAAB49627	GTP-binding protei
426	5	16.7	161	21	AAAG19950	Arabinidopsis thalia	499	5	16.7	197	22	AAAB60072	Human protein SEQ
427	5	16.7	161	21	AAAG46009	Arabinidopsis thalia	500	5	16.7	197	22	AAAB79505	Hericidially activ
428	5	16.7	161	22	ABG02697	Novel human diagno	501	5	16.7	200	23	ABBB92146	Arabinidopsis thalia
429	5	16.7	162	13	AAAR21140	Polypeptide encode	502	5	16.7	205	21	AAAG53203	Arabinidopsis thalia
430	5	16.7	162	21	AAAB54164	Human pancreatic c	503	5	16.7	206	21	AAAB42001	Human OREF ORF765
431	5	16.7	162	22	ABBB5554	Drosophila melanog	504	5	16.7	208	20	AAAG3590	Human TNRL3 protei
432	5	16.7	162	23	ABBS3404	Lactococcus lactis	505	5	16.7	209	22	ABG29051	Novel human diagno
433	5	16.7	164	20	AAV06951	E. chaffeensis OMP-	506	5	16.7	209	22	AAU23499	Novel human enzyme
434	5	16.7	164	22	ABBB1833	Drosophila melanog	507	5	16.7	210	23	ABP25506	Streptococcus poly
435	5	16.7	165	21	AAAG23658	Arabinidopsis thalia	508	5	16.7	211	20	AAAG3591	Mouse TNRL3 protei
436	5	16.7	166	21	AAAG51578	Arabinidopsis thalia	509	5	16.7	211	20	AAAY95918	Human Goodpasture
437	5	16.7	166	22	AAU23140	Novel human enzyme	510	5	16.7	212	22	AAU45973	Novel human diagno
438	5	16.7	166	22	AAU23688	Novel human enzyme	511	5	16.7	212	22	ABG17051	Novel human diagno
439	5	16.7	167	20	AAV38569	Nisseria meningit	512	5	16.7	215	21	AAV79494	Elmeria tenella 25
440	5	16.7	167	22	ABBB97760	Human secretory po	513	5	16.7	215	21	AAAG06197	Arabinidopsis thalia
441	5	16.7	169	22	AAE06783	Human dual -specifi	514	5	16.7	215	21	AAAG12970	Arabinidopsis thalia
442	5	16.7	169	22	AAAB6439	Human MAP-kinase p	515	5	16.7	215	21	AAAG47703	Arabinidopsis thalia
443	5	16.7	170	19	AAAB64200	Human calcineurin.	516	5	16.7	215	22	AAU37468	Staphylococcus aur
444	5	16.7	170	21	AAAB09977	Human CNRII protei	517	5	16.7	215	23	ABP42892	Human ovarian anti
445	5	16.7	170	21	AAAB09978	Human HCNB protein	518	5	16.7	215	23	AAAB1956	Hericidially activ
446	5	16.7	170	22	AAO14411	Calcineurin B subu	519	5	16.7	217	20	AAV13950	HIV gp120 v3 loop.
447	5	16.7	170	22	ABBB60493	Drosophila melanog	520	5	16.7	218	16	AAAR79164	Partial sequence o
448	5	16.7	171	22	AAU17530	Novel signal trans	521	5	16.7	218	20	AAAY44172	Human type IV coll

522	5	16.7	218	21	AAV56784	Human alpha3 type	595	5	16.7	261	20	AAV38934	Neisseria gonorrhoea
523	5	16.7	218	22	AAE09484	Human alpha-3 chain	596	5	16.7	261	21	AAAB42113	Human ORFX ORF1877
524	5	16.7	219	21	AAAG29766	Arabidopsis thaliana	597	5	16.7	261	21	AAAG24387	Human ORFX ORF1877
525	5	16.7	219	21	AAAG57933	Arabidopsis thaliana	598	5	16.7	262	21	AAAB25330	Bacillus subtilis
526	5	16.7	220	22	AAU02911	Angiotensin convertase	599	5	16.7	263	22	AAU51961	Propionibacterium
527	5	16.7	220	23	AAU10477	Mouse pancreatic p	600	5	16.7	264	22	AAAM0361	Human polypeptide
528	5	16.7	221	21	AAAG17322	Arabidopsis thaliana	601	5	16.7	264	22	AAAB94175	Human protein sequ
529	5	16.7	223	22	AAAB97131	Corynebacterium di	602	5	16.7	265	22	ABAG16595	Novel human diagno
530	5	16.7	224	22	AAAB69844	Nfgr protein. Unu	603	5	16.7	265	23	ABP29956	Streptococcus epi
531	5	16.7	225	19	AAW47524	Mus musculus tumo	604	5	16.7	265	23	ABP29951	Streptococcus poly
532	5	16.7	225	21	AAAB07527	Amino acid sequenc	605	5	16.7	266	20	AAV31993	Type IV collagen N
533	5	16.7	225	21	AAAG34261	Arabidopsis thaliana	606	5	16.7	268	20	AAV37555	Human alpha3(IV)N
534	5	16.7	225	22	ABAG16600	Novel human diagno	607	5	16.7	268	22	ABAG15429	Novel human diagno
535	5	16.7	226	12	AAAR13901	Nitrite hydratase	608	5	16.7	269	22	ABAG10707	Novel human diagno
536	5	16.7	226	21	AAAG14311	Arabidopsis thaliana	609	5	16.7	270	18	AAW29816	Mammalian AMPK bet
537	5	16.7	226	21	AAAG53230	Arabidopsis thaliana	610	5	16.7	270	23	ABP28995	Streptococcus poly
538	5	16.7	226	22	AAAG80788	C diphtheriae DtxR	611	5	16.7	271	22	ABP28995	Streptococcus poly
539	5	16.7	228	21	AAAG15809	Arabidopsis thaliana	612	5	16.7	271	22	ABP28995	Streptococcus poly
540	5	16.7	229	10	AAAP94802	Pb1r HIV fusion p	613	5	16.7	273	21	AAAG46007	Human microglobulin
541	5	16.7	229	11	AAAR04494	HIV fusion protein	614	5	16.7	273	22	AAU03499	Arabidopsis thaliana
542	5	16.7	229	21	AAAG44499	Arabidopsis thaliana	615	5	16.7	274	21	AAAG51580	Arabidopsis thaliana
543	5	16.7	229	21	AAAG53227	Arabidopsis thaliana	616	5	16.7	274	23	ABP25484	Streptococcus poly
544	5	16.7	230	23	ABBS4940	Lactococcus lactis	617	5	16.7	275	20	AAV06956	E. chaffeensis OMP-
545	5	16.7	231	22	ABG24191	Novel human diagno	618	5	16.7	275	21	AAAG05586	E. chaffeensis OMP-
546	5	16.7	234	9	AAAP81143	Sequence of fusion	619	5	16.7	275	23	AAAG05586	E. chaffeensis OMP-
547	5	16.7	234	10	AAAP2910	HTLV-III fusion pr	620	5	16.7	276	21	AAAG24386	Arabidopsis thaliana
548	5	16.7	234	14	AAAR1943	Recombinant PBI fu	621	5	16.7	276	21	AAV96466	Partial soybean 4-
549	5	16.7	234	22	AAAG1970	S. epidermidis ope	622	5	16.7	276	22	ABG05655	Novel human diagno
550	5	16.7	235	21	AAAG15808	Arabidopsis thaliana	623	5	16.7	276	22	ABG05655	Novel human diagno
551	5	16.7	236	21	AAAG29765	Arabidopsis thaliana	624	5	16.7	278	22	AAE04866	Corynebacterium gl
552	5	16.7	236	23	ABP26002	Streptococcus poly	625	5	16.7	279	23	ABG21482	Novel human diagno
553	5	16.7	236	23	ABP26003	Streptococcus poly	626	5	16.7	279	23	ABBB4768	DNA polymerase III
554	5	16.7	238	22	ABBB1493	Drosophila melanog	627	5	16.7	280	20	AAV06949	E. chaffeensis OMP-
555	5	16.7	239	22	AAAG98390	Escherichia coli p	628	5	16.7	282	19	AAW60150	M. vaccae truncate
556	5	16.7	240	21	AAAG53202	Arabidopsis thaliana	629	5	16.7	282	20	AAV14897	M. vaccae truncate
557	5	16.7	243	21	AAAG29764	Arabidopsis thaliana	630	5	16.7	282	23	ABR73503	M vaccae GV-38A tr
558	5	16.7	244	23	AAU075595	Human type IV coll	631	5	16.7	284	19	AAW47525	Human secreted pro
559	5	16.7	245	21	AAAG19949	Arabidopsis thaliana	632	5	16.7	284	22	ABG15436	Novel human diagno
560	5	16.7	245	21	AAAG46024	Arabidopsis thaliana	633	5	16.7	284	23	ABBB6193	Human kidney injur
561	5	16.7	245	21	AAV67942	Human type IV coll	634	5	16.7	284	23	AAAT7831	Herpes simplex vir
562	5	16.7	245	23	AAU075589	Human type IV coll	635	5	16.7	285	20	AAV06957	E. chaffeensis OMP-
563	5	16.7	246	20	AAV60028	Human endometrium	636	5	16.7	285	21	AAAB25373	P. pinus radiata cell
564	5	16.7	246	22	AAAB47009	Wild type E. coli	637	5	16.7	285	21	AAV76141	Human secreted pro
565	5	16.7	248	21	AAV68779	Amino acid sequenc	638	5	16.7	285	23	ABBB9378	Human polypeptide
566	5	16.7	249	19	AAV29745	TNF related endothe	639	5	16.7	285	23	AAU73408	Ehrlichia chaffe
567	5	16.7	249	20	AAV09369	Human tumour necro	640	5	16.7	286	21	AAAG51575	Arabidopsis thaliana
568	5	16.7	249	21	AAAB07526	Amino acid sequenc	641	5	16.7	286	22	AAU023781	Novel human enzyme
569	5	16.7	249	22	AAAB05338	Human TREPA (TNF r	642	5	16.7	288	21	AAAB41910	Human ORFX ORF1674
570	5	16.7	249	23	AAAB00891	Human TREPA (TNF r	643	5	16.7	288	21	AAAG46023	Arabidopsis thaliana
571	5	16.7	249	23	ABP45149	Human BLYS binding	644	5	16.7	289	20	AAV06966	E. canis p30-6 pro
572	5	16.7	250	23	AAU061129	Human PRO207 poly	645	5	16.7	289	21	AAAB44522	Virulence gene pro
573	5	16.7	250	21	AAV97158	Human alpha-2-macr	646	5	16.7	289	21	AAV95553	Chlamydia pneumonia
574	5	16.7	251	10	AAAP1333	Peptide expressed	647	5	16.7	289	22	ABBB4137	Drosophila melanog
575	5	16.7	251	15	AAAR1885	BioC gene prod. in	648	5	16.7	291	23	ABP29525	Streptococcus poly
576	5	16.7	251	23	AAV92372	p75-NTR (neurotrop	649	5	16.7	291	23	AAU73409	Ehrlichia chaffe
577	5	16.7	251	23	ABP45677	Human BLYS binding	650	5	16.7	292	20	AAV06967	E. canis p30-7 pro
578	5	16.7	251	23	ABBB6194	Human kidney injur	651	5	16.7	292	22	ABG29598	Novel human diagno
579	5	16.7	252	21	AAAG47392	Arabidopsis thaliana	652	5	16.7	293	20	AAV06965	E. canis p30-5 pro
580	5	16.7	252	22	AAAG33024	C glutamicum prote	653	5	16.7	293	21	AAAG50529	Arabidopsis thaliana
581	5	16.7	254	24	AAU075598	Human type IV coll	654	5	16.7	293	23	AAU096115	Ehrlichia canis p2
582	5	16.7	255	20	AAW87729	Rupescris stem pit	655	5	16.7	293	23	AAU73412	Ehrlichia chaffe
583	5	16.7	256	21	AAAG31623	Arabidopsis thaliana	656	5	16.7	294	22	ABBS52678	Escherichia coli p
584	5	16.7	257	21	AAAG50530	Arabidopsis thaliana	657	5	16.7	294	22	ABBS52678	Escherichia coli p
585	5	16.7	257	21	AAAG51577	Arabidopsis thaliana	658	5	16.7	294	22	ABBS3033	Escherichia coli p
586	5	16.7	257	23	ABBB9570	Listeria monocytog	659	5	16.7	295	20	AAV06953	E. chaffeensis OMP-
587	5	16.7	258	21	AAAG15807	Arabidopsis thaliana	660	5	16.7	295	23	AAU073404	Ehrlichia chaffe
588	5	16.7	258	23	ABBB92143	Streptococcus poly	661	5	16.7	297	8	AAV71198	Sequence of HTLV-I
589	5	16.7	258	23	ABBB92143	Herbicidally activ	662	5	16.7	297	20	AAV14916	Extended amino act
590	5	16.7	259	20	AAV38931	Neisseria meningit	663	5	16.7	297	21	AAAG44498	Arabidopsis thaliana
591	5	16.7	259	23	ABBB92476	Herbicidally activ	664	5	16.7	297	21	AAAG53226	Arabidopsis thaliana
592	5	16.7	260	21	AAAG44355	Arabidopsis thaliana	665	5	16.7	297	22	ABBB68339	Drosophila melanog
593	5	16.7	260	21	AAAG51581	Arabidopsis thaliana	666	5	16.7	297	23	ABBB73522	M vaccae GV-38A tr
594	5	16.7	260	23	ABBB92475	Herbicidally activ	667	5	16.7	298	23	AAU73410	Ehrlichia chaffe

668	5	16.7	300	19	AAV66084	S. pneumoniae deri	741	5	16.7	337	20	AAV30158	Rat dorsal root re
669	5	16.7	300	21	AAAG8272	Arabidopsis thalia	742	5	16.7	337	21	AAAG6830	Arabidopsis thalia
670	5	16.7	300	22	AAAG8163	S. epidermidis ope	743	5	16.7	337	21	AAAG6870	Arabidopsis thalia
671	5	16.7	301	21	AAAG24811	Arabidopsis thalia	744	5	16.7	338	21	AAAG07164	Arabidopsis thalia
672	5	16.7	301	21	AAV57945	Human transmembran	745	5	16.7	338	22	AAU37864	Streptococcus pneu
673	5	16.7	301	22	AAV35049	Human polypeptide	746	5	16.7	338	22	AAU38052	Streptococcus pneu
674	5	16.7	301	22	AAU12254	Human PR04343 poly	747	5	16.7	339	21	AAAG05228	Arabidopsis thalia
675	5	16.7	301	22	AAV53638	Lactococcus lactis	748	5	16.7	339	22	AAU09964	Human cytidine dea
676	5	16.7	302	22	AAV76876	Human lung tumour	749	5	16.7	339	22	AAU30263	Novel human secret
677	5	16.7	302	23	AAU95531	US515 lung tumour	750	5	16.7	340	21	AAAG14354	Arabidopsis thalia
678	5	16.7	304	20	AAV35465	Chlamydia pneumoni	751	5	16.7	341	22	ABG31567	Novel human diagno
679	5	16.7	304	21	AAAG46022	Arabidopsis thalia	752	5	16.7	341	23	AAV50720	Haemophilus paraga
680	5	16.7	305	23	AAU75568	Human dendritic ce	753	5	16.7	341	23	AAV50723	Haemophilus paraga
681	5	16.7	306	20	AAV30814	Human secreted pro	754	5	16.7	341	23	AAV50726	Haemophilus paraga
682	5	16.7	307	22	ABV64255	Drosophila melanog	755	5	16.7	342	14	AAV39489	Human apoAIV mutel
683	5	16.7	309	12	AAV31175	NS1.81-Rfdeletag	756	5	16.7	342	14	AAV39491	Human apoAIV mutel
684	5	16.7	310	23	ABV54948	Lactococcus lactis	757	5	16.7	342	14	AAV39496	Human apoAIV mutel
685	5	16.7	312	23	AAO17807	H influenzae BVH-N	758	5	16.7	342	14	AAV39498	Human apoAIV mutel
686	5	16.7	313	21	AAAG38271	Arabidopsis thalia	759	5	16.7	344	21	AAAG12969	Arabidopsis thalia
687	5	16.7	313	22	AAAG82332	S. epidermidis ope	760	5	16.7	344	21	AAAG14353	Neisseria meningit
688	5	16.7	314	22	AAV41627	Human ORFX ORF1391	761	5	16.7	344	21	AAV5045	Haemophilus paraga
689	5	16.7	314	22	AAV63384	Putative P. abyssi	762	5	16.7	344	23	AAV50716	Haemophilus paraga
690	5	16.7	314	22	AAV40835	Human polypeptide	763	5	16.7	344	23	AAV50717	Haemophilus paraga
691	5	16.7	315	22	AAV70792	N. meningitidis bacter	764	5	16.7	344	23	AAV50718	Haemophilus paraga
692	5	16.7	316	21	AAV25566	Eucalyptus grandis	765	5	16.7	344	23	AAV50719	Haemophilus paraga
693	5	16.7	316	21	AAV50528	Arabidopsis thalia	766	5	16.7	344	23	AAV50721	Haemophilus paraga
694	5	16.7	317	19	AAV72043	HSV-2 strain SBS C	767	5	16.7	344	23	AAV50722	Haemophilus paraga
695	5	16.7	317	22	ABG30146	Novel human diagno	768	5	16.7	344	23	AAV50724	Haemophilus paraga
696	5	16.7	317	22	ABV39692	Staphylococcus epi	769	5	16.7	344	23	AAV50725	Haemophilus paraga
697	5	16.7	319	8	AAV71199	Sequence of HTLV-I	770	5	16.7	344	23	AAV50727	Haemophilus paraga
698	5	16.7	319	11	AAV07945	NS181RfAuch plasm	771	5	16.7	346	14	AAV39493	Human apoAIV mutel
699	5	16.7	319	12	AAV13176	Drosophila melanog	772	5	16.7	347	22	AAV39494	Chenopodium album
700	5	16.7	320	22	ABV67007	Arabidopsis thalia	773	5	16.7	347	23	AAV39495	Chenopodium album
701	5	16.7	320	22	ABV70584	HIV gp120 protein	774	5	16.7	347	23	AAV39496	Lactococcus lactis
702	5	16.7	321	20	AAV14209	HIV gp120 protein	775	5	16.7	348	23	ABV47897	Listeria monocytog
703	5	16.7	321	20	AAV14205	Escherichia coli b	776	5	16.7	349	22	AAV58340	Propionibacterium
704	5	16.7	322	20	AAV13601	Amino acid sequenc	777	5	16.7	349	22	AAV58340	Human olfactory re
705	5	16.7	323	21	AAV07758	Arabidopsis thalia	778	5	16.7	349	22	AAV58340	Human membrane or
706	5	16.7	323	22	ABG18805	Novel human diagno	779	5	16.7	350	21	AAV70278	Recombinant vaccin
707	5	16.7	324	21	AAV34260	Arabidopsis thalia	780	5	16.7	350	22	ABG01198	Novel human diagno
708	5	16.7	325	21	AAV13668	Arabidopsis thalia	781	5	16.7	351	21	AAV44591	Virulence gene pro
709	5	16.7	325	21	AAV38270	Arabidopsis thalia	782	5	16.7	352	22	AAV2753	Human protein sequ
710	5	16.7	326	21	AAV07166	Arabidopsis thalia	783	5	16.7	353	10	AAV2015	HIV portion of fus
711	5	16.7	326	21	AAV46832	Arabidopsis thalia	784	5	16.7	353	10	AAV3538	AT portion of fus
712	5	16.7	327	12	AAV13177	NS1.81-RfAuch + (	785	5	16.7	353	21	AAV43352	Arabidopsis thalia
713	5	16.7	327	12	AAV09096	Tomato polygalactu	786	5	16.7	353	21	AAV43352	Arabidopsis thalia
714	5	16.7	328	14	AAV39484	Human apoAIV mutel	787	5	16.7	357	22	AAV2759	Human protein sequ
715	5	16.7	328	14	AAV34259	Arabidopsis thalia	788	5	16.7	363	14	AAV39478	Human apoAIV mutel
716	5	16.7	328	23	ABV38384	Staphylococcus epi	789	5	16.7	363	14	AAV39479	Human apoAIV mutel
717	5	16.7	329	22	ABG16057	Novel human diagno	790	5	16.7	363	22	ABG28654	Novel human diagno
718	5	16.7	329	22	ABG16224	Novel human diagno	791	5	16.7	365	21	AAV5565	Arabidopsis thalia
719	5	16.7	330	21	AAV07165	Arabidopsis thalia	792	5	16.7	365	22	AAV4163	Human protein sequ
720	5	16.7	330	21	AAV46831	Arabidopsis thalia	793	5	16.7	366	22	ABV57854	Drosophila melanog
721	5	16.7	330	21	AAV46871	Arabidopsis thalia	794	5	16.7	366	22	ABV57854	Novel human diagno
722	5	16.7	330	22	ABG13035	Novel human diagno	795	5	16.7	366	22	AAV6277	Putative P. abyssi
723	5	16.7	330	22	AAV79895	Corynebacterium gl	796	5	16.7	367	21	AAV2566	Arabidopsis thalia
724	5	16.7	330	23	ABV33378	Herbicidally activ	797	5	16.7	369	22	AAV1904	C glutamicum prote
725	5	16.7	330	23	ABV33379	Herbicidally activ	798	5	16.7	371	21	AAV5585	Arabidopsis thalia
726	5	16.7	331	22	AAV91206	C glutamicum prote	799	5	16.7	371	22	ABG13741	Novel human diagno
727	5	16.7	332	21	AAV05229	Arabidopsis thalia	800	5	16.7	371	22	ABG14939	Novel human diagno
728	5	16.7	332	22	ABG24501	Novel human diagno	801	5	16.7	372	22	ABG04187	Novel human diagno
729	5	16.7	333	14	AAV39488	Human apoAIV mutel	802	5	16.7	373	14	AAV39486	Human apoAIV mutel
730	5	16.7	333	14	AAV39490	Human apoAIV mutel	803	5	16.7	374	22	AAV61161	Streptomyces livid
731	5	16.7	333	14	AAV39495	Human apoAIV mutel	804	5	16.7	375	23	ABV6286	Streptococcus poly
732	5	16.7	333	14	AAV39497	Human apoAIV mutel	805	5	16.7	377	14	AAV39483	Human apoAIV mutel
733	5	16.7	334	12	ABG08757	Novel human diagno	806	5	16.7	377	14	AAV39480	Human apoAIV mutel
734	5	16.7	335	12	AAV13178	NS1.81 (NANP) 4RLFAU	807	5	16.7	377	14	AAV39499	Human apoAIV mutel
735	5	16.7	335	12	AAV13179	NS1.81 (NANP) 4RLFAU	808	5	16.7	377	14	AAV39500	Human apoAIV mutel
736	5	16.7	335	22	AAV90576	C glutamicum prote	809	5	16.7	377	14	AAV39502	Human apoAIV mutel
737	5	16.7	336	22	ABV52466	Escherichia coli p	810	5	16.7	377	14	AAV39501	Human apoAIV mutel
738	5	16.7	337	14	AAV39485	Human apoAIV mutel	811	5	16.7	377	14	AAV5242	Human apoAIV mutel
739	5	16.7	337	14	AAV39492	Human apoAIV mutel	812	5	16.7	377	14	AAV5243	Human apoAIV mutel
740	5	16.7	337	14	AAV39494	Human apoAIV mutel	813	5	16.7	377	14	AAV5244	Human apoAIV mutel



814	5	16.7	377	22	AA893692	Human protein sequ	887	5	16.7	411	17	AAW00276	Cyanobacterial lye
815	5	16.7	377	23	AAOI4939	Human UFD2-associa	888	5	16.7	412	7	AAE0416	CS protein of mala
816	5	16.7	378	21	AA613816	Arabidopsis thalia	889	5	16.7	412	9	AAE0835	Sequence encoded b
817	5	16.7	378	22	AA631676	Arabidopsis thalia	890	5	16.7	413	23	AB893471	Hericicidally activ
818	5	16.7	378	22	AB865454	Drosophila melanog	891	5	16.7	414	23	AB849456	Listeria monocytog
819	5	16.7	379	23	AAOI7802	H influenzae BYN-N	892	5	16.7	416	14	AA844434	Subtilisin-like se
820	5	16.7	381	20	AAI78570	Neisseria meningit	893	5	16.7	416	22	AB858589	Drosophila melano
821	5	16.7	381	21	AAI75046	Neisseria meningit	894	5	16.7	417	16	AA870649	Mouse azospermia
822	5	16.7	382	20	AAI758571	Neisseria gonorrhoe	895	5	16.7	419	22	AAU40408	Propionibacterium
823	5	16.7	382	21	AAI75044	Neisseria gonorrhoe	896	5	16.7	419	22	AAU40408	Escherichia coli p
824	5	16.7	383	22	AAAI1928	Human polypeptide	897	5	16.7	420	23	AB854041	Lactococcus lactis
825	5	16.7	384	22	AAAI1781	Arabidopsis thalia	898	5	16.7	421	22	AA870801	S cerevisiae apopt
826	5	16.7	384	21	AA648096	Arabidopsis thalia	899	5	16.7	421	22	AA868332	Amino acid sequenc
827	5	16.7	384	22	AB871205	Drosophila melanog	900	5	16.7	423	10	AA892013	HIV portion of HTL
828	5	16.7	386	21	AA825527	Pinus radiata cell	901	5	16.7	423	10	AA893536	Sequence of HIV po
829	5	16.7	386	21	AA613815	Arabidopsis thalia	902	5	16.7	423	21	AA605563	Arabidopsis thalia
830	5	16.7	386	22	AA604290	Human gene 11 enco	903	5	16.7	423	21	AA614141	Arabidopsis thalia
831	5	16.7	386	23	AB893375	Herbicidally activ	904	5	16.7	423	21	AA612447	Arabidopsis thalia
832	5	16.7	388	22	AA851345	Mouse HS-glycoprot	905	5	16.7	424	14	AA837796	RTS protein. Synt
833	5	16.7	389	21	AA605564	Arabidopsis thalia	906	5	16.7	424	16	AA866202	Complete human bon
834	5	16.7	389	22	AB865000	Drosophila melanog	907	5	16.7	424	21	AAI92029	Human bone morphog
835	5	16.7	389	23	AB889418	Human polypeptide	908	5	16.7	425	21	AAI92370	p75-NTR (neurotrop
836	5	16.7	393	21	AA605584	Arabidopsis thalia	909	5	16.7	425	22	AB829201	Novel human diagno
837	5	16.7	393	21	AA613814	Arabidopsis thalia	910	5	16.7	425	23	AA821671	Rat neurotrophic r
838	5	16.7	394	15	AA63601	MAP-kinase-phospha	911	5	16.7	426	21	AA620523	Arabidopsis thalia
839	5	16.7	394	16	AA878635	Partial MAP kinase	912	5	16.7	427	20	AAI33483	Human neutrophil r
840	5	16.7	394	22	AA876875	Human lung tumour	913	5	16.7	427	22	AA819582	Novel human diagno
841	5	16.7	394	22	AA876878	Human lung tumour	914	5	16.7	427	22	AA836699	Human tumour necro
842	5	16.7	394	22	AA657448	Amino acid sequenc	915	5	16.7	427	22	AA850894	Human neurotrophic
843	5	16.7	394	22	AA657627	Amino acid sequenc	916	5	16.7	427	23	AA821670	Human neurotrophic
844	5	16.7	394	23	AAU85530	SV40 lung tumour p	917	5	16.7	427	23	AB892637	Hericicidally activ
845	5	16.7	394	23	AAU85533	L515 lung tumour p	918	5	16.7	427	23	AB853860	Lactococcus lactis
846	5	16.7	396	18	AAW34218	Streptomyces hydro	919	5	16.7	428	23	AAU92983	Beta-primeverosida
847	5	16.7	396	19	AAW55819	Streptomyces roseo	920	5	16.7	429	21	AA826811	Listeria monocytog
848	5	16.7	396	22	AA890664	Human secreted pro	921	5	16.7	429	23	AB848663	Novel human diagno
849	5	16.7	396	23	AAU10860	Human apolipoprote	922	5	16.7	431	22	AB806289	Novel human diagno
850	5	16.7	396	23	AAU10861	Human apolipoprote	923	5	16.7	431	22	AA895196	Human protein sequ
851	5	16.7	396	23	AAU10862	Human apolipoprote	924	5	16.7	432	22	AA890924	C glutamicum prote
852	5	16.7	396	23	AAU10863	Human apolipoprote	925	5	16.7	434	22	AA890945	Human sbg6804SPAR
853	5	16.7	396	23	AAU10864	Human apolipoprote	926	5	16.7	434	23	AA817534	Human protein modi
854	5	16.7	396	23	AAU10865	Human apolipoprote	927	5	16.7	435	21	AA819384	Amino acid sequenc
855	5	16.7	396	23	AAU10866	Human apolipoprote	928	5	16.7	435	21	AA895050	Arabidopsis thalia
856	5	16.7	396	23	AAU10867	Human apolipoprote	929	5	16.7	435	22	AA839070	Human polypeptide
857	5	16.7	396	23	AAU10868	Human apolipoprote	930	5	16.7	435	23	AB848846	Listeria monocytog
858	5	16.7	396	23	AAU10869	Human apolipoprote	931	5	16.7	436	21	AA611674	Arabidopsis thalia
859	5	16.7	398	21	AA829612	Cat flea HMT ANON/	932	5	16.7	436	22	AA618343	Novel human diagno
860	5	16.7	398	22	AB859946	Drosophila melanog	933	5	16.7	437	23	AB878795	Human NOV1 protein
861	5	16.7	398	23	AB853914	Lactococcus lactis	934	5	16.7	440	23	AA823386	Human intracellular
862	5	16.7	399	21	AA819396	Amino acid sequenc	935	5	16.7	440	23	AB847784	Protein #30 relate
863	5	16.7	399	22	AB815380	Novel human diagno	936	5	16.7	440	23	AB826921	Streptococcus poly
864	5	16.7	399	22	AA893874	Human polypeptide,	937	5	16.7	441	23	AAU09884	Novel human secret
865	5	16.7	399	22	AA804208	Human gene 11 enco	938	5	16.7	443	22	AA805813	Human small cell 1
866	5	16.7	399	22	AA804230	Human gene 11 enco	939	5	16.7	445	16	AA872451	Brevibacterium fla
867	5	16.7	399	23	AB864487	Human albumin fusi	940	5	16.7	445	18	AAW23285	Brevibacterium lac
868	5	16.7	399	23	AB864488	Human albumin fusi	941	5	16.7	445	18	AAW06588	B. lactofermentum
869	5	16.7	401	22	AB803895	Novel human diagno	942	5	16.7	445	19	AAW68150	Diaminopimelate de
870	5	16.7	401	22	AB810083	Novel human diagno	943	5	16.7	445	19	AAW69552	Brevibacterium lac
871	5	16.7	401	22	AB813038	Novel human diagno	944	5	16.7	445	19	AAW47399	B. lactofermentum
872	5	16.7	401	23	AAU85617	Recombinant protei	945	5	16.7	445	22	AA866811	Human zinc finger
873	5	16.7	402	21	AA631675	Arabidopsis thalia	946	5	16.7	445	22	AAW78643	Human protein SEQ
874	5	16.7	403	21	AA620623	Arabidopsis thalia	947	5	16.7	445	22	AA893197	C glutamicum prote
875	5	16.7	405	23	AB840414	Staphylococcus epi	948	5	16.7	445	22	AA804862	Corynebacterium gl
876	5	16.7	405	20	AAW82680	Streptomyces sp. t	949	5	16.7	445	22	AA804863	Corynebacterium gl
877	5	16.7	406	22	AAW39645	Human polypeptide	950	5	16.7	445	22	AA894631	Human protein sequ
878	5	16.7	406	23	AB861840	Prostate cancer-as	951	5	16.7	445	23	AB889471	Human polypeptide
879	5	16.7	407	23	AB870686	Drosophila melanog	952	5	16.7	446	21	AA854714	Arabidopsis thalia
880	5	16.7	408	22	AB819521	Novel human diagno	953	5	16.7	447	20	AAI35554	C. pneumoniae prot
881	5	16.7	408	22	AA809452	Human sbg6804SPAR	954	5	16.7	448	23	AA854422	Lactococcus lactis
882	5	16.7	408	22	AAW41431	Human polypeptide	955	5	16.7	449	21	AAI96476	Soybean 4-lydroxyp
883	5	16.7	408	23	AB808246	Human SPARC-homolo	956	5	16.7	449	22	AA892396	C glutamicum prote
884	5	16.7	409	22	AAU67750	Propionibacterium	957	5	16.7	449	22	AA878891	C. glutamicum SRT
885	5	16.7	411	9	AA831144	Sequence encoded b	958	5	16.7	450	23	AAW50772	Rice ubiquitin lig
886	5	16.7	411	16	AA876980	Lycopene cyclase.	959	5	16.7	451	16	AA863088	Breast tumor kinas

960	5	16.7	451	22	ABG02799	Novel human diagno
961	5	16.7	452	22	ABB68662	Drosophila melanog
962	5	16.7	452	22	ABB59808	TutC protein. Tha
963	5	16.7	453	21	AA642565	Arabidopsis thalia
964	5	16.7	455	19	AAW60147	M. vaccae antigen
965	5	16.7	455	20	AAV14894	Amino acid sequenc
966	5	16.7	455	21	AAAB26983	Human NGFR p75. H
967	5	16.7	455	23	ABB73500	M vaccae GV-38A pa
968	5	16.7	458	20	AAV38932	Neisseria meningit
969	5	16.7	458	20	AAV38933	Neisseria meningit
970	5	16.7	458	20	AAV38935	Neisseria gonorrhoe
971	5	16.7	458	23	ABP30342	Streptococcus poly
972	5	16.7	458	23	AAU72978	Neisseria meningit
973	5	16.7	459	21	AAV14140	Arabidopsis thalia
974	5	16.7	459	21	AAV32446	Arabidopsis thalia
975	5	16.7	459	22	AAV79656	Corynebacterium gl
976	5	16.7	460	22	AAU46872	Propionibacterium
977	5	16.7	465	21	AAV31991	Arabidopsis thalia
978	5	16.7	465	22	AAV95369	Human protein sequ
979	5	16.7	465	23	ABG65646	Human breast speci
980	5	16.7	466	21	AAV32073	Arabidopsis thalia
981	5	16.7	466	23	ABP28095	Streptococcus poly
982	5	16.7	466	23	ABV54231	Lactococcus lactis
983	5	16.7	467	23	ABP28094	Streptococcus poly
984	5	16.7	469	19	AAW37061	Sequence of bovine
985	5	16.7	469	19	AAW37061	HIV-1 breakthrough
986	5	16.7	469	20	AAV44171	Bovine papillomavi
987	5	16.7	470	22	ABV55574	Drosophila melanog
988	5	16.7	470	22	ABG16644	Novel human diagno
989	5	16.7	471	16	AAV79163	Partial sequence o
990	5	16.7	471	20	AAV44171	Bovine type IV col
991	5	16.7	471	20	AAV14914	Extended amino aci
992	5	16.7	471	21	AAV50499	Arabidopsis thalia
993	5	16.7	471	21	AAV56783	Bovine alpha3 type
994	5	16.7	471	22	AAE09483	M vaccae alpha-3 cha
995	5	16.7	471	22	ABV73520	M vaccae GV-38A pr
996	5	16.7	472	6	AAV50288	Sequence encoded b
997	5	16.7	473	23	AAE23641	White mustard myr3
998	5	16.7	474	22	ABV63130	Drosophila melanog
999	5	16.7	475	22	ABV58603	Drosophila melanog
1000	5	16.7	475	23	ABP30445	Streptococcus poly

## ALIGNMENTS

RESULT 1  
ID AAW53252 standard; Protein; 30 AA.

AAW53252;  
30-JUL-1998 (first entry)

Candida albicans allergen - antigenic protein SEQ ID NO:2.

Candida albicans; vaccine; allergen; antigenic protein; fungal antigen;  
immune response; infection; insoluble.

Candida albicans.

Key Location/Qualifiers

Misc-difference 26 /note= "any amino acid"

Misc-difference 29 /note= "any amino acid"

Misc-difference 30 /note= "any amino acid"

W09803990-A1.

12-MAR-1998.

XX

PF 29-AUG-1997; 97WO-JP03041.  
XX 31-MAR-1997; 97JP-0099775.  
PR 04-SEP-1996; 96JP-0255400.  
XX (TAKI ) TAKARA SHUZO CO LTD.  
XX Endo M, Kato I, Mizutani S, Takesako K;  
PI WPI; 1998-193553/17.  
XX Fungal antigens comprising insoluble fraction of fungal cells -  
PT useful for, e.g. stimulating immune response and treatment and  
PT diagnosis of fungal infection(s)  
PS Claim 26; Page 75; 108pp; Japanese.  
XX The present sequence represents an active vaccine component or allergen  
CC derived from Candida albicans, which is an antigenic protein. The  
CC present invention describes fungal antigens, comprising the insoluble  
CC fraction of fungal cells having completely/partially removed cell walls.  
CC Also described are nucleic acids encoding the antigens and a method for  
CC producing the antigens. The antigens can be used for preparing  
CC therapeutic compositions for stimulating immune response, e.g. as a  
CC vaccine. They can be used for treatment of fungal infections, treatment  
CC and prevention of allergies and diagnosis of fungal infections in  
CC vertebrates. The vaccines are not live, and have low toxicity.  
XX Sequence 30 AA;  
SQ  
Query Match 90.0%; Score 27; DB 19; Length 30;  
Best Local Similarity 100.0%; Pred. No. 3.1e-22;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KYSLPELDYERSATPRYSQINELIXYT 28  
Db 1 KYSLPELDYERSATPRYSQINELIXYT 28  
RESULT 2  
ID ABB05612 standard; Protein; 233 AA.  
XX ABB05612;  
XX 24-APR-2002 (first entry)  
XX S. cerevisiae manganese superoxide dismutase protein SEQ ID NO:31.  
XX Aspergillus oryzae; mutant; modified; reduced transcription; hormone;  
XX reduced translation; reduced secretion; receptor; antibody; reporter;  
XX enzyme; lipase; manganese superoxide dismutase.  
XX Saccharomyces cerevisiae.  
XX Synthetic.  
XX US6323002-B1.  
XX 27-NOV-2001.  
XX 25-JUN-1999; 99US-0339972.  
XX 12-SEP-1997; 97US-0928692.  
XX 13-SEP-1996; 96US-0713312.  
XX (NOVO ) NOVOZYMES BIOTECH INC.  
XX Brody H, Yaver DS, Lamsa M, Hansen K;  
PI WPI; 2002-163017/21.  
XX Producing a polypeptide using a cell for reducing the production of the  
PT polypeptide, comprises inserting DNA into the genome of the cell at a

PT position not within the polypeptide coding sequence or a regulatory  
 XX Sequence -  
 PS Example 17; Column 125-128; 129pp; English.  
 CC The present invention describes a method for producing a polypeptide (PI)  
 CC comprising cultivating a mutant cell whose parent cell comprises a DNA  
 CC sequence encoding PI, by introducing a nucleic acid construct into the  
 CC genome of the parent cell at a locus not within the PI sequence, so  
 CC that PI transcription, translation or secretion is reduced, and  
 CC recovering PI. The method is used to produce a polypeptide, such as a  
 CC recombinant or heterologous hormone, hormone variant, receptor, antibody,  
 CC reporter or enzyme, particularly an oxidoreductase, transferase,  
 CC hydrolase, lyase, isomerase or ligase. The present sequence represents  
 CC a Saccharomyces cerevisiae manganese superoxide dismutase protein which  
 CC shares significant identity with the mutant Aspergillus oryzae DB571058  
 CC protein, which is used in an example from the present invention.  
 CC Sequence 233 AA;  
 Query Match 33.3%; Score 10; DB 23; Length 233;  
 Best Local Similarity 100.0%; Pred. No. 0.0053;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CY 15 EPIYSGQINE 24  
 DB 41 EPIYSGQINE 50  
 RESULT 3  
 ID AAG14176 standard; Protein; 407 AA.  
 AC AAG14176;  
 XX 17-OCT-2000 (first entry)  
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 13938.  
 XX  
 KM Protein identification; signal transduction pathway; metabolic pathway;  
 KM hybridisation assay; genetic mapping; gene expression control; promoter;  
 KM termination sequence.  
 XX Arabidopsis thaliana.  
 OS  
 PN BPI033405-A2.  
 XX  
 XX 06-SEP-2000.  
 XX  
 XX 25-FEB-2000; 2000EP-0301439.  
 XX  
 XX 25-FEB-1999; 99US-0121825.  
 PR 05-MAR-1999; 99US-0123180.  
 PR 09-MAR-1999; 99US-0123548.  
 PR 23-MAR-1999; 99US-0125788.  
 PR 25-MAR-1999; 99US-0126264.  
 PR 29-MAR-1999; 99US-0126785.  
 PR 01-APR-1999; 99US-0127462.  
 PR 06-APR-1999; 99US-0128234.  
 PR 08-APR-1999; 99US-0128714.  
 PR 16-APR-1999; 99US-0129845.  
 PR 19-APR-1999; 99US-0130077.  
 PR 21-APR-1999; 99US-0130449.  
 PR 23-APR-1999; 99US-0130510.  
 PR 23-APR-1999; 99US-0130891.  
 PR 28-APR-1999; 99US-0131449.  
 PR 30-APR-1999; 99US-0132048.  
 PR 04-MAY-1999; 99US-0132407.  
 PR 05-MAY-1999; 99US-0132484.  
 PR 06-MAY-1999; 99US-0132485.  
 PR 06-MAY-1999; 99US-0132486.  
 PR 06-MAY-1999; 99US-0132487.  
 PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.  
 PR 14-MAY-1999; 99US-0134218.  
 PR 14-MAY-1999; 99US-0134219.  
 PR 14-MAY-1999; 99US-0134221.  
 PR 14-MAY-1999; 99US-0134370.  
 PR 18-MAY-1999; 99US-0134768.  
 PR 19-MAY-1999; 99US-0134941.  
 PR 20-MAY-1999; 99US-0135124.  
 PR 21-MAY-1999; 99US-0135353.  
 PR 24-MAY-1999; 99US-0135629.  
 PR 25-MAY-1999; 99US-0136021.  
 PR 27-MAY-1999; 99US-0136382.  
 PR 28-MAY-1999; 99US-0136782.  
 PR 01-JUN-1999; 99US-0137222.  
 PR 03-JUN-1999; 99US-0137528.  
 PR 04-JUN-1999; 99US-0137502.  
 PR 07-JUN-1999; 99US-0137724.  
 PR 08-JUN-1999; 99US-0138094.  
 PR 10-JUN-1999; 99US-0138540.  
 PR 10-JUN-1999; 99US-0138847.  
 PR 14-JUN-1999; 99US-0139119.  
 PR 16-JUN-1999; 99US-0139451.  
 PR 16-JUN-1999; 99US-0139452.  
 PR 17-JUN-1999; 99US-0139453.  
 PR 17-JUN-1999; 99US-0139454.  
 PR 18-JUN-1999; 99US-0139455.  
 PR 18-JUN-1999; 99US-0139456.  
 PR 18-JUN-1999; 99US-0139457.  
 PR 18-JUN-1999; 99US-0139458.  
 PR 18-JUN-1999; 99US-0139459.  
 PR 18-JUN-1999; 99US-0139460.  
 PR 18-JUN-1999; 99US-0139461.  
 PR 18-JUN-1999; 99US-0139462.  
 PR 18-JUN-1999; 99US-0139463.  
 PR 18-JUN-1999; 99US-0139750.  
 PR 18-JUN-1999; 99US-0139763.  
 PR 21-JUN-1999; 99US-0139817.  
 PR 22-JUN-1999; 99US-0139899.  
 PR 23-JUN-1999; 99US-0140353.  
 PR 24-JUN-1999; 99US-0140354.  
 PR 24-JUN-1999; 99US-0140635.  
 PR 28-JUN-1999; 99US-0140823.  
 PR 29-JUN-1999; 99US-0140991.  
 PR 30-JUN-1999; 99US-0141287.  
 PR 01-JUL-1999; 99US-0141842.  
 PR 01-JUL-1999; 99US-0142154.  
 PR 02-JUL-1999; 99US-0142055.  
 PR 06-JUL-1999; 99US-0142390.  
 PR 08-JUL-1999; 99US-0142803.  
 PR 09-JUL-1999; 99US-0142920.  
 PR 12-JUL-1999; 99US-0142977.  
 PR 13-JUL-1999; 99US-0143542.  
 PR 14-JUL-1999; 99US-0143624.  
 PR 15-JUL-1999; 99US-0144005.  
 PR 16-JUL-1999; 99US-0144085.  
 PR 16-JUL-1999; 99US-0144086.  
 PR 19-JUL-1999; 99US-0144325.  
 PR 19-JUL-1999; 99US-0144331.  
 PR 19-JUL-1999; 99US-0144332.  
 PR 19-JUL-1999; 99US-0144333.  
 PR 19-JUL-1999; 99US-0144334.  
 PR 19-JUL-1999; 99US-0144335.  
 PR 20-JUL-1999; 99US-0144352.  
 PR 20-JUL-1999; 99US-0144632.  
 PR 20-JUL-1999; 99US-0144884.  
 PR 21-JUL-1999; 99US-0144814.  
 PR 21-JUL-1999; 99US-0145086.  
 PR 21-JUL-1999; 99US-0145088.  
 PR 22-JUL-1999; 99US-0145085.  
 PR 22-JUL-1999; 99US-0145087.  
 PR 22-JUL-1999; 99US-0145089.  
 PR 22-JUL-1999; 99US-0145192.  
 PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0145386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158233.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.

PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.  
PR 29-OCT-1999; 99US-0162143.

Query Match 23.3%; Score 7; DB 21; Length 407;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 SATEPYI 18  
Db 131 SATEPYI 137

RESULT 4  
AAG14175  
ID AAG14175 standard; Protein; 512 AA.  
XX  
AC AAG14175;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 13937.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
OS  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139753.  
PR 18-JUN-1999; 99US-0139817.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144864.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148655.  
PR 13-AUG-1999; 99US-0148664.  
PR 16-AUG-1999; 99US-0149358.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149910.  
PR 25-AUG-1999; 99US-0150565.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 18-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160771.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.

PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.  
Query Match 23.3%; Score 7; DB 21; Length 512;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 12 SATPEYI 18  
DB 226 SATPEYI 242  
RESULT 5  
AAG14174  
ID AAG14174 standard; Protein; 521 AA.  
AC AAG14174;  
XX 17-OCT-2000 (first entry)  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 13936.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hydridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0121180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0128645.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132485.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139452.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139753.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139859.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146388.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.

```

PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149375.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 14-OCT-1999; 99US-0159644.
PR 16-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160761.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Query Match 23.3%; Score 7; DB 21; Length 521;

```

Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 12 SATPEPY 18
Db 245 SATPEPY 251

```

```

RESULT 6
ABB08465
ID ABB08465 standard; Peptide; 7 AA.

```

```

XX ABB08465;

```

```

DT 29-JUL-2002 (first entry)

```

```

DE Mouse tumour specific antigenic related protein #1.

```

```

XX Proton pump inhibitor; tumour-specific antigenic peptide;

```

```

KW cyostatic; tumour; mouse.

```

```

OS Mus sp.

```

```

PN JP2001286284-A.

```

```

PD 16-OCT-2001.

```

```

PF 05-APR-2000; 2000JP-0103966.

```

```

XX 05-APR-2000; 2000JP-0103966.

```

```

XX (SATO/) SATO N.

```

```

PA (SUZU/) SUZUKI N.

```

```

PA (YAMA/) YAMAGUCHI M.

```

```

DR WPI; 2002-134186/18.

```

```

XX Gene diagnostic agent and gene treating agent for tumours comprises

```

```

PT using tumour-specific antigen and proton pump inhibitor as antitumour

```

```

PT agent

```

```

XX Example 2; Page 30; 80pp; Japanese.

```

```

CC This invention relates to a diagnostic agent or treating agent for

```

```

CC tumours by using a tumour specific antigen and a proton pump

```

```

CC inhibitor. The method is cyostatic in its action and the

```

```

CC tumour-specific antigen peptide is useful for treating and preventing

```

```

CC cancers. This amino acid sequence is related to the tumour specific

```

```

XX Sequence 7 AA;

```

```

Qy 11 FSATBP 16

```

```

Db 1 FSATBP 6

```

```

RESULT 7
ABB08476
ID ABB08476 standard; Peptide; 7 AA.

```

```

XX ABB08476;

```

```

DT 29-JUL-2002 (first entry)

```

```

DE Peptide related to tumour specific antigen.

```

```

XX Proton pump inhibitor; tumour-specific antigenic peptide;

```

```

KW cyostatic; tumour.

```

XX Unidentified.  
OS  
XX  
XX JP2001286284-A.  
PN  
XX  
XX 16-OCT-2001.  
PD  
XX  
XX 05-APR-2000; 2000JP-0103966.  
PF  
XX  
XX 05-APR-2000; 2000JP-0103966.  
PR  
XX  
XX 05-APR-2000; 2000JP-0103966.  
PS  
XX  
XX (SATO/) SATO N.  
PA  
XX (SUZU/) SUZUKI N.  
XX (YAMA/) YAMAGUCHI M.  
DR  
XX  
XX WPI; 2002-134186/18.  
PT  
XX  
XX Gene diagnostic agent and gene creating agent for tumours comprises  
PT using tumour-specific antigen and proton pump inhibitor as antitumour  
PT agent -  
XX  
XX Claim 23; Page 4; 80pp; Japanese.  
XX  
XX  
XX This invention relates to a diagnostic agent or treating agent for  
CC tumours by using a tumour specific antigen and a proton pump  
CC inhibitor. The method is cytostatic in its action and the  
CC tumour-specific antigen peptide is useful for treating and preventing  
CC cancers. This amino acid sequence is related to the tumour specific  
CC antigen and proton pump of the invention.  
XX  
XX  
XX Sequence 7 AA;  
SQ  
XX  
XX Query Match 20.0%; Score 6; DB 23; Length 7;  
Best Local Similarity 100.0%; Pred. No. 7.8e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 11 FSATEP 16  
DB 1 FSATEP 6  
XX  
XX  
XX RESULT 8  
AB08466  
ID ABB08466 standard; Peptide; 9 AA.  
XX  
XX  
XX ABB08466;  
AC  
XX  
XX 29-JUL-2002 (first entry)  
DT  
XX  
XX Mouse tumour specific antigenic related protein #2.  
XX  
XX Proton pump inhibitor; tumour-specific antigenic peptide;  
XX cyostatic; tumour; mouse.  
XX  
XX Mus sp.  
OS  
XX  
XX JP2001286284-A.  
PN  
XX  
XX 16-OCT-2001.  
PD  
XX  
XX 05-APR-2000; 2000JP-0103966.  
PF  
XX  
XX 05-APR-2000; 2000JP-0103966.  
PR  
XX  
XX 05-APR-2000; 2000JP-0103966.  
PS  
XX  
XX (SATO/) SATO N.  
PA  
XX (SUZU/) SUZUKI N.  
XX (YAMA/) YAMAGUCHI M.  
DR  
XX  
XX WPI; 2002-134186/18.  
PT  
XX  
XX Gene diagnostic agent and gene creating agent for tumours comprises  
PT using tumour-specific antigen and proton pump inhibitor as antitumour  
PT agent -

XX Claim 25; Page 30; 80pp; Japanese.  
PS  
XX  
XX  
XX This invention relates to a diagnostic agent or treating agent for  
CC tumours by using a tumour specific antigen and a proton pump  
CC inhibitor. The method is cytostatic in its action and the  
CC tumour-specific antigen peptide is useful for treating and preventing  
CC cancers. This amino acid sequence is related to the tumour specific  
CC antigen and proton pump of the invention.  
XX  
XX  
XX Sequence 9 AA;  
SQ  
XX  
XX Query Match 20.0%; Score 6; DB 23; Length 9;  
Best Local Similarity 100.0%; Pred. No. 7.8e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 11 FSATEP 16  
DB 2 FSATEP 7  
XX  
XX  
XX RESULT 9  
AB08467  
ID ABB08467 standard; Peptide; 10 AA.  
XX  
XX  
XX ABB08467;  
AC  
XX  
XX 29-JUL-2002 (first entry)  
DT  
XX  
XX Mouse tumour specific antigenic related protein #3.  
XX  
XX Proton pump inhibitor; tumour-specific antigenic peptide;  
XX cyostatic; tumour; mouse.  
XX  
XX Mus sp.  
OS  
XX  
XX JP2001286284-A.  
PN  
XX  
XX 16-OCT-2001.  
PD  
XX  
XX 05-APR-2000; 2000JP-0103966.  
PF  
XX  
XX 05-APR-2000; 2000JP-0103966.  
PR  
XX  
XX 05-APR-2000; 2000JP-0103966.  
PS  
XX  
XX (SATO/) SATO N.  
PA  
XX (SUZU/) SUZUKI N.  
XX (YAMA/) YAMAGUCHI M.  
DR  
XX  
XX WPI; 2002-134186/18.  
PT  
XX  
XX Gene diagnostic agent and gene creating agent for tumours comprises  
PT using tumour-specific antigen and proton pump inhibitor as antitumour  
PT agent -  
XX  
XX Claim 25; Page 30; 80pp; Japanese.  
XX  
XX  
XX This invention relates to a diagnostic agent or treating agent for  
CC tumours by using a tumour specific antigen and a proton pump  
CC inhibitor. The method is cytostatic in its action and the  
CC tumour-specific antigen peptide is useful for treating and preventing  
CC cancers. This amino acid sequence is related to the tumour specific  
CC antigen and proton pump of the invention.  
XX  
XX  
XX Sequence 10 AA;  
SQ  
XX  
XX Query Match 20.0%; Score 6; DB 23; Length 10;  
Best Local Similarity 100.0%; Pred. No. 5.5;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 11 FSATEP 16  
DB 3 FSATEP 8



RESULT 10  
 ABG66263  
 ID ABG66263 standard; Peptide; 19 AA.  
 XX  
 AC ABG66263;  
 XX  
 DT 29-AUG-2002 (first entry)  
 XX  
 DE IGE Fcepsilon RI binding peptide from phage displayed g3 library IGE154.  
 XX  
 KW IGE receptor; immunoglobulin; FcepsilonRI; antagonist; phage display;  
 KW protein co-ordinate data; IGE-mediated disease; allergic rhinitis;  
 KW asthma; allergic asthma; atopic dermatitis; urticaria-angioedema;  
 KW parasitic infection; IGE myeloma; immune-related disorder;  
 KW inflammatory disorder; diabetes mellitus; reperfusion injury; stroke;  
 KW IGE-mediated gastrointestinal inflammatory disease; burn;  
 KW immune rejection of graft; myocardial infarction; atherosclerosis;  
 KW acute lung injury; haemorrhagic shock; septic shock;  
 KW acute tubular necrosis; endometriosis; degenerative joint disease;  
 KW pancreatitis.  
 XX  
 OS Synthetic.  
 XX  
 PN WO200226781-A2.  
 XX  
 PD 04-APR-2002.  
 XX  
 PF 26-SEP-2001; 2001WO-US30289.  
 XX  
 PR 26-SEP-2000; 2000US-235353P.  
 PR 23-MAR-2001; 2001US-278540P.  
 XX  
 XX (GETH ) GENENTECH INC.  
 XX  
 PI Lowman HB, Reynolds ME, Nakamura GR, Starovassnik MA;  
 XX  
 DR WPI; 2002-444016/47.  
 XX  
 PT A peptide useful for treating a IGE-mediated disease or disorder in a  
 PT host e.g. allergic rhinitis, asthma, which competes with immunoglobulin  
 PT E for binding to high affinity IGE receptor in an in vitro assay -  
 XX  
 XX Example 12; Page 120; 328pp; English.  
 XX  
 CC The invention relates to a peptide which competes with immunoglobulin  
 CC (Ig) E 134 comprising a sequence (S1), for binding the high affinity IGE  
 CC receptor (FcepsilonRI) in an in vitro assay and having a formula given in  
 CC the specification. Also included are a fusion protein comprising the  
 CC peptide, a pharmaceutical composition (C) comprising the peptide,  
 CC designing a compound that mimics the three-dimensional surface  
 CC structure of the peptide, a compound with a solvent accessible surface  
 CC that mimics the solvent accessible surface defined by the side chains of  
 CC residues (R) Pro4, Phe6, Pro16, Cys3, Cys7, Cys15 and Cys19 of IGE134,  
 CC a peptide with structural coordinates as given in the specification,  
 CC selecting a peptide mimetic which binds to FcepsilonRI and blocks  
 CC binding of IGE and a peptide mimetic which mimics the coordinates of  
 CC IGE134 residues (R). (C) is useful for inhibiting the binding of IGE to  
 CC high affinity IGE receptor (FcepsilonRI). Peptides of the formula given  
 CC in the specification are useful for inhibiting the binding of an IGE to  
 CC molecule which blocks the interaction of IGE with high affinity IGE  
 CC receptor. The peptide is also useful for inhibiting the activation of  
 CC high affinity IGE receptor. The peptide is useful for treating an IGE-  
 CC mediated disease or disorder in a host. (C) is useful in research,  
 CC diagnostic, therapeutic and prophylactic methods. The peptide is also  
 CC useful for inhibiting IGE-mediated or associated processes such as IGE-  
 CC dependent activation and degranulation of mast cells and basophils, as  
 CC well as consequent release of inflammatory mediators such as histamine.  
 CC (C) is useful for treating allergic rhinitis, asthma (e.g. allergic  
 CC asthma), atopic dermatitis, urticaria-angioedema, parasitic infection,  
 CC IGE myeloma, immune-related disorders, inflammatory disorders, diabetes  
 CC mellitus, IGE-mediated gastrointestinal inflammatory disease, immune

CC rejection of grafts, reperfusion injury, stroke, myocardial infarction,  
 CC atherosclerosis, acute lung injury, haemorrhagic shock, burn, septic  
 CC shock, acute tubular necrosis, endometriosis, degenerative joint disease  
 CC and pancreatitis. The present sequence is a peptide of the invention  
 CC expressed from a phage display library.  
 XX  
 SQ Sequence 19 AA;  
 XX  
 Query Match 20.0%; Score 6; DB 23; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 10;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 5 PELDYE 10  
 DQ 7 PELDYE 12  
 XX  
 RESULT 11  
 ABB08468  
 ID ABB08468 standard; Protein; 20 AA.  
 XX  
 AC ABB08468;  
 XX  
 DT 29-JUL-2002 (first entry)  
 XX  
 DE Mouse tumour specific antigenic related protein #4.  
 XX  
 KW Proton pump inhibitor; tumour-specific antigenic peptide;  
 KW cyrostatic; tumour; mouse.  
 XX  
 OS Mus sp.  
 XX  
 PN JP2001286284-A.  
 XX  
 PD 16-OCT-2001.  
 XX  
 PF 05-APR-2000; 2000JP-0103966.  
 XX  
 PR 05-APR-2000; 2000JP-0103966.  
 XX  
 PA (SATO/) SATO N.  
 PA (SUZUKI) SUZUKI N.  
 PA (YAMA/) YAMAGUCHI M.  
 XX  
 DR WPI; 2002-134186/18.  
 XX  
 PT Gene diagnostic agent and gene treating agent for tumours comprises  
 PT using tumour-specific antigen and proton pump inhibitor as antitumour  
 PT agent -  
 XX  
 PS Claim 25; Page 30; 80pp; Japanese.  
 XX  
 CC This invention relates to a diagnostic agent or treating agent for  
 CC tumours by using a tumour specific antigen and a proton pump  
 CC inhibitor. The method is cytostatic in its action and the  
 CC tumour-specific antigen peptide is useful for treating and preventing  
 CC cancers. This amino acid sequence is related to the tumour specific  
 CC antigen and proton pump of the invention.  
 XX  
 SQ Sequence 20 AA;  
 XX  
 Query Match 20.0%; Score 6; DB 23; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 11;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 11 FSATEP 16  
 DQ 4 FSATEP 9  
 XX  
 RESULT 12  
 AAB78878  
 ID AAB78878 standard; Protein; 44 AA.

XX AAB78878;  
AC  
XX  
DT 30-APR-2001 (first entry)  
XX  
DE C. glutamicum SRT protein sequence SEQ ID NO:16.  
XX  
KW Corynebacterium glutamicum; stress; resistance; tolerance; SRT;  
KW fine chemical production; organic acid; proteiogenic amino acid;  
KW nonproteogenic amino acid; purine base; pyrimidine base; nucleoside;  
KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;  
KW carboxylate; aromatic compound; vitamin; cofactor; polypeptide; enzyme;  
KW evolutionary study; environmental hazard; fermentation.  
XX  
OS Corynebacterium glutamicum.  
XX  
PN WO200100804-A2.  
XX  
PD 04-JAN-2001.  
XX  
PF 23-JUN-2000; 2000WO-1B00922.  
XX  
PR 25-JUN-1999; 99US-0141031.  
PR 01-JUL-1999; 99DE-1030429.  
PR 01-JUL-1999; 99US-0142692.  
PR 08-JUL-1999; 99DE-1031413.  
PR 08-JUL-1999; 99DE-1031457.  
PR 08-JUL-1999; 99DE-1031541.  
PR 09-JUL-1999; 99DE-1032209.  
PR 09-JUL-1999; 99DE-1032230.  
PR 14-JUL-1999; 99DE-1032914.  
PR 27-AUG-1999; 99DE-1040764.  
PR 27-AUG-1999; 99US-0151214.  
PR 31-AUG-1999; 99DE-1041382.  
XX  
PA (BADI ) BASF AG.  
XX  
PI Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberman G, Lee H,  
PI Kim H;  
XX  
XX WPI; 2001-061972/07.  
DR N-PSDB; AAF70991.  
XX  
XX New isolated Corynebacterium glutamicum nucleic acid encoding a stress,  
PT tolerance or resistance protein, for production or modulation of  
PT production of fine chemicals, such as, e.g. amino acids, lipide,  
PT carbohydrates, or enzymes -  
XX  
XX  
PS Claim 20; Page 138; 526pp; English.  
XX  
XX AAF70984 to AAF71133 encode the Corynebacterium glutamicum stress,  
CC resistance, and tolerance (SRT) proteins given in AAB78871 to AAB79020.  
CC The C. glutamicum SRT genes (I) can be used in vectors (II) for  
CC expression in host cells and production of fine chemicals, such as, an  
CC organic acid, a proteiogenic or nonproteogenic amino acid (preferred),  
CC a purine or pyrimidine base, a nucleoside, a nucleotide, a lipid, a  
CC saturated or unsaturated fatty acid, a diol, a carboxylate, an aromatic  
CC compound, a vitamin, a cofactor, a polypeptide, or an enzyme. The fine  
CC chemical production can be modulated. The presence of (I) or the SRT  
CC or proteins (III) encoded by them are used for diagnosing the presence  
CC or activity of Corynebacterium diptheriae. (I), (II), (III) and host  
CC cells containing them can be used to map the genomes of organisms related  
CC to C. glutamicum, to identify and localise C. glutamicum sequences of  
CC interest, in evolutionary studies, in determination of SRT protein  
CC regions required for function, in modulating the SRT protein activity,  
CC and in modulating the activity of an SRT pathway. (II) are used to permit  
CC C. glutamicum to survive in an environment that is normally  
CC environmentally or chemically hazardous to it. (I) and protein molecules  
CC encoded by it increase the survival of C. glutamicum to chemical and  
CC environmental hazards and provide a means for continued growth and  
CC multiplication in large scale fermentative growth conditions. By  
CC increasing the growth rate or maintaining a normal growth rate in poor or  
CC toxic conditions, the yield, production and/or efficiency or production

CC of fine chemicals from a culture may be increased.  
XX  
XX  
SQ Sequence 44 AA;  
XX  
XX Query Match 20.0%; Score 6; DB 22; Length 44;  
XX Best Local Similarity 100.0%; Pred. No. 23;  
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 4 LPELDY 9  
DB 6 LPELDY 11  
XX  
XX RESULT 13  
XX ABB36416  
XX ID ABB36416 standard; Peptide; 50 AA.  
XX  
XX ABB36416;  
XX  
XX 04-FEB-2002 (first entry)  
XX  
XX Peptide #3922 encoded by human foetal liver single exon probe.  
XX  
XX Human; foetal liver; gene expression; single exon nucleic acid probe.  
XX  
XX Homo sapiens.  
XX  
XX WO200157277-A2.  
XX  
XX 09-AUG-2001.  
XX  
XX 30-JAN-2001; 2001WO-US00669.  
XX  
XX 04-FEB-2000; 2000US-0180312.  
XX 26-MAY-2000; 2000US-0207456.  
XX 30-JUN-2000; 2000US-0608408.  
XX 03-AUG-2000; 2000US-0632366.  
XX 21-SEP-2000; 2000US-0234687.  
XX 27-SEP-2000; 2000US-0236359.  
XX 04-OCT-2000; 2000GB-0024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
XX WPI; 2001-483447/52.  
XX  
XX Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human fetal liver -  
PT  
XX  
XX Claim 27; SEQ ID NO 29051; 639pp + sequence listing; English.  
XX  
XX The invention relates to a single exon nucleic acid probe for  
CC measuring human gene expression in a sample derived from human foetal  
CC liver. The single exon nucleic acid probes may be used for predicting,  
CC measuring and displaying gene expression in samples derived from human  
CC foetal liver. The present sequence is a peptide encoded by a single exon  
CC nucleic acid probe of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO. [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences).  
XX  
SQ Sequence 50 AA;  
XX  
XX Query Match 20.0%; Score 6; DB 22; Length 50;  
XX Best Local Similarity 100.0%; Pred. No. 26;  
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 12 SATPEY 17  
DB 10 SATPEY 15

```

RESULT 14
AAM57176
ID AAM57176 standard; Protein; 50 AA.
XX
XX
AC AAM57176;
XX
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 29281.
XX
XX Human brain expressed exon; gene expression analysis; probe;
KM microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer.
XX
XX Homo sapiens.
OS
PN WO200157275-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00667.
XX
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
XX
XX Example 4; SEQ ID NO: 29281; 650bp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system,
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention.
XX
XX
SQ Sequence 50 AA;
XX
XX Query Match 20.0%; Score 6; DB 22; Length 50;
XX Best Local Similarity 100.0%; Pred. No. 26;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 12 SATPEY 17
XX |||||
DB 10 SATPEY 15
XX

RESULT 15
AAM69579
ID AAM69579 standard; Protein; 50 AA.
XX
XX AAM69579;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human bone marrow expressed probe encoded protein SEQ ID NO: 29885.
DE Human bone marrow expressed exon; gene expression analysis; probe;
XX microarray; cancer; leukaemia; lymphoma; myeloma.
XX
XX

```

```

OS Homo sapiens.
XX
XX WO200157276-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00668.
XX
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488900/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
XX Example 4; SEQ ID NO: 29885; 658bp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention.
XX
XX
SQ Sequence 50 AA;
XX
XX Query Match 20.0%; Score 6; DB 22; Length 50;
XX Best Local Similarity 100.0%; Pred. No. 26;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 12 SATPEY 17
XX |||||
DB 10 SATPEY 15
XX

RESULT 16
AAM17397
ID AAM17397 standard; Protein; 50 AA.
XX
XX AAM17397;
XX
XX 12-OCT-2001 (first entry)
XX
XX Peptide #3831 encoded by probe for measuring cervical gene expression.
DE Probe; human; microarray; gene expression; cervical epithelial cell;
XX cervical cancer.
XX
XX Homo sapiens.
OS
XX
XX WO200157278-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00670.
XX
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX

```

XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-488901/53.  
XX Human genome-derived single exon nucleic acid probes useful for  
XX analyzing gene expression in human cervical epithelial cells -  
XX  
XX Claim 27; SEQ ID No 22223; 487bp; English.  
XX  
XX The present invention relates to human single exon nucleic acid probes  
XX (SENPs: see A110068-A1128459). The present sequence is a peptide encoded  
XX by one such probe. The SENPs are derived from human HeLa cells. The SENPs  
XX can be used to produce a single exon microarray, which can be used for  
XX measuring human gene expression in a sample derived from human cervical  
XX epithelial cells. By measuring gene expression, the probes are therefore  
XX useful in grading and/or staging of diseases of the cervix, notably  
XX cervical cancer.  
XX Note: The sequence data for this patent did not form part of the printed  
XX specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 50 AA;  
XX  
XX Query March 20.0%; Score 6; DB 22; Length 50;  
XX Best Local Similarity 100.0%; Pred. No. 26;  
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX 12 SATPEY 17  
XX 10 SATPEY 15  
XX  
XX RESULT 17  
XX AAM29917  
XX ID AAM29917 standard; Protein; 50 AA.  
XX  
XX AAM29917;  
XX  
XX 17-OCT-2001 (first entry)  
XX  
XX Peptide #3954 encoded by probe for measuring placental gene expression.  
XX  
XX Probe: microarray; human; placenta; antenatal diagnosis;  
XX genetic disorder.  
XX  
XX Homo sapiens.  
XX  
XX WO200157272-A2.  
XX  
XX 09-AUG-2001.  
XX  
XX 30-JAN-2001; 2001WO-US00663.  
XX  
XX 04-FEB-2000; 2000US-0180312.  
XX 26-MAY-2000; 2000US-0207456.  
XX 30-JUN-2000; 2000US-0608408.  
XX 03-AUG-2000; 2000US-0632366.  
XX 21-SEP-2000; 2000US-0234687.  
XX 27-SEP-2000; 2000US-0236359.  
XX 04-OCT-2000; 2000GB-0024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-488901/53.  
XX Human genome-derived single exon nucleic acid probes useful for  
XX analyzing gene expression in human placenta -  
XX

PS Claim 27; SEQ ID No 30186; 654bp; English.  
XX  
XX The present invention relates to single exon nucleic acid probes (SENPs:  
XX see A113115-A157546). The present sequence is a peptide encoded by one  
XX such probe. The probes are useful for producing a microarray for  
XX predicting, measuring and displaying gene expression in samples derived  
XX from human placenta. The probes are useful for antenatal diagnosis of  
XX human genetic disorders.  
XX  
XX Sequence 50 AA;  
XX  
XX Query March 20.0%; Score 6; DB 22; Length 50;  
XX Best Local Similarity 100.0%; Pred. No. 26;  
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX 12 SATPEY 17  
XX 10 SATPEY 15  
XX  
XX RESULT 18  
XX ABG39201  
XX ID ABG39201 standard; Peptide; 50 AA.  
XX  
XX ABG39201;  
XX  
XX 19-AUG-2002 (first entry)  
XX  
XX Human peptide encoded by genome-derived single exon probe SEQ ID 28866.  
XX  
XX Human; single exon probe; asthma; lung cancer; COPD; ILD;  
XX chronic obstructive pulmonary disease; interstitial lung disease;  
XX familial idiopathic pulmonary fibrosis; neurofibromatosis;  
XX tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;  
XX Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;  
XX pulmonary histiocytosis; lymphangioleiomyomatosis; Karsagen syndrome;  
XX pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;  
XX primary ciliary dyskinesia; pulmonary hypertension;  
XX hyaline membrane disease.  
XX  
XX Homo sapiens.  
XX  
XX WO200186003-A2.  
XX  
XX 15-NOV-2001.  
XX  
XX 30-JAN-2001; 2001WO-US00665.  
XX  
XX 04-FEB-2000; 2000US-180312P.  
XX 26-MAY-2000; 2000US-207456P.  
XX 30-JUN-2000; 2000US-0608408.  
XX 03-AUG-2000; 2000US-0632366.  
XX 21-SEP-2000; 2000US-234687P.  
XX 27-SEP-2000; 2000US-236359P.  
XX 04-OCT-2000; 2000GB-0024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2002-114183/15.  
XX  
XX Spatially-addressable set of single exon nucleic acid probes, used to  
XX measure gene expression in human lung samples -  
XX  
XX Claim 27; SEQ ID No 28866; 634bp; English.  
XX  
XX The invention relates to a spatially-addressable set of single exon  
XX nucleic acid probes for measuring gene expression in a sample derived  
XX from human lung comprising single exon nucleic acid probes having one of  
XX 12614 nucleic acid sequences mentioned in the specification, or their  
XX complements or the 1287 open reading frames derived from the 12614  
XX probes. Also included are a microarray comprising the novel set of

CC probes; the novel set of probes which hybridize at high stringency to a  
CC nucleic acid expressed in the human lung; measuring gene expression in a  
CC sample derived from human lung, comprising (a) contacting the array with  
CC a collection of detectably labeled nucleic acids derived from human lung  
CC mRNA, and (b) measuring the label detectably bound to each probe of  
CC the array; identifying exons in a eukaryotic genome, comprising  
CC (a) algorithmically predicting at least one exon from genomic sequences  
CC of the eukaryote; and (b) detecting specific hybridization of detectably  
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
CC having a fragment identical to the predicted exon, the probe is included  
CC in the above mentioned microarray; assigning exons to a single gene,  
CC comprising (a) identifying exons from genomic sequence by the method  
CC above and (b) measuring the expression of each of the exons in several  
CC tissues and/or cell types using hybridization to a single exon  
CC microarrays having a probe with the exon, where a common pattern of  
CC expression of the exons in the tissues and/or cell types indicates that  
CC the exons should be assigned to a single gene; a peptide comprising one  
CC of 12011 sequences, mentioned in the specification, or encoded by the  
CC probes/open reading frames (ORF). The probes are used for gene  
CC expression analysis, and for identifying exons in a gene, particularly  
CC using human lung derived mRNA and for the study of lung diseases  
CC such as asthma, lung cancer, chronic obstructive pulmonary disease  
CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary  
CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,  
CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary  
CC hemangioendothelioma, pulmonary histiocytosis, lymphangioleiomyomatosis,  
CC pulmonary alveolar proteinosis, Karsenger syndrome, fibrocystic  
CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension  
CC and hyaline membrane disease. The present sequence is a peptide/protein  
CC encoded by a single exon probe of the invention.  
CC Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic  
CC format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

SQ Sequence 50 AA;

Query Match 20.0%; Score 6; DB 23; Length 50;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 SATPEY 17  
| | | | |  
Db 10 SATPEY 15

RESULT 19

ID AAV58434 standard; Protein; 52 AA.  
AAV58434;

XX AAV58434;

XX 27-FEB-2002 (first entry)

XX Propionibacterium acnes immunogenic protein #19330.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
XX uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
XX inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
XX dermatological; osteopathic; neuroprotectant.

XX Propionibacterium acnes.

XX WO200181561-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-US12865.

XX 21-APR-2000; 2000US-199047P.

XX 02-JUN-2000; 2000US-208841P.

XX 07-JUL-2000; 2000US-216747P.

PA (CORI-) CORIXA CORP.

XX Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;

XX L'maisonmeuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.

XX N-PSDB; AAS59591.

XX Propionibacterium acnes polypeptides and nucleic acids useful for

XX treating acne vulgaris -

XX Example 1; SEQ ID No 19629; 1069bp; English.

CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic  
CC polypeptides. The proteins and their associated DNA sequences are used in  
CC the treatment, prevention and diagnosis of medical conditions caused by  
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.  
CC P. acnes is also involved in infections of bone, joints and the central  
CC nervous system, however it is particularly involved in the inflammatory  
CC lesions associated with acne vulgaris. A method for detecting the  
CC presence or absence of P. acnes in a patient comprises contacting a  
CC sample with a binding agent that binds to the proteins of the invention  
CC and determining the amount of bound protein in the sample. The  
CC polypeptides may be used as antigens in the production of antibodies  
CC specific for P. acnes proteins. These antibodies can be used to  
CC downregulate expression and activity of P. acnes polypeptides and  
CC therefore treat P. acnes infections. The antibodies may also be used as  
CC diagnostic agents for determining P. acnes presence, for example, by  
CC enzyme linked immunosorbent assay (ELISA).  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

SQ Sequence 52 AA;

Query Match 20.0%; Score 6; DB 22; Length 52;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 SATPEY 17  
| | | | |  
Db 9 SATPEY 14

RESULT 20

ID AAU65746 standard; Protein; 52 AA.  
AAU65746;

XX AAU65746;

XX 27-FEB-2002 (first entry)

XX Propionibacterium acnes immunogenic protein #26642.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
XX uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
XX inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
XX dermatological; osteopathic; neuroprotectant.

XX Propionibacterium acnes.

XX WO200181561-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-US12865.

XX 21-APR-2000; 2000US-199047P.

XX 02-JUN-2000; 2000US-208841P.

XX 07-JUL-2000; 2000US-216747P.



PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 20-OCT-2000; 2000US-0242221.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246539.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249287.  
PR 17-NOV-2000; 2000US-0249289.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250391.  
PR 01-DEC-2000; 2000US-0251160.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
PI  
XX WPI; 2001-541565/60.  
XX N-PSDB; ABA11135.  
XX  
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating nervous system  
PT cancers and metastases -

XX Claim 11; SEQ ID NO 3466; 1701pp + Sequence Listing; English.  
PS  
XX The invention relates to novel genes (ABA11004-ABA21534) and proteins  
CC (ABA14678-ABH18001) useful for preventing, treating or ameliorating  
CC medical conditions e.g. by protein or gene therapy. The genes are  
CC isolated from a range of human tissues disclosed in the specification.  
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful  
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast  
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone  
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
CC hemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;  
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
CC and parasitic infections.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at [http://wipo.int/pub/published\\_pct\\_sequences](http://wipo.int/pub/published_pct_sequences).  
XX

SO Sequence 60 AA;

Query Match 20.0%; Score 6; DB 22; Length 60;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 DYERFA 13  
Db 50 DYERFA 55

#### RESULT 22

ABR08458  
ID ABR08458 standard; Protein; 62 AA.

XX ABR08458;

DT 29-JUL-2002 (first entry)

XX Mouse tumour specific antigenic peptide #2.

DE Proton pump inhibitor; tumour-specific antigenic peptide;

KM cytosolic; tumour; murine.

XX Mus musculus.

PN JP2001286284-A.

PD 16-OCT-2001.

PF 05-APR-2000; 2000JP-0103966.

XX 05-APR-2000; 2000JP-0103966.

PA (SATO/) SATO N.

PA (SUZU/) SUZUKI N.

PA (YAMA/) YAMAGUCHI M.

XX WPI; 2002-134186/18.

DR N-PSDB; ABA99214.

XX Gene diagnostic agent and gene treating agent for tumours comprises  
PT using tumour-specific antigen and proton pump inhibitor as antitumour  
PT agent -

XX Claim 24; Page 49; 80pp; Japanese.

XX This invention relates to a diagnostic agent or treating agent for  
CC tumours by using a tumour specific antigen and a proton pump  
CC inhibitor. The method is cytosolic in its action and the  
CC tumour-specific antigen peptide is useful for treating and preventing

CC cancers. This sequence is a mouse tumour-specific  
XX antigenic peptide.  
SQ Sequence 62 AA;  
Query Match 20.0%; Score 6; DB 23; Length 62;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 11 PSATEP 16  
Db 49 PSATEP 54  
RESULT 23  
AAW75089  
ID AAW75089 standard; Protein; 81 AA.  
XX AAW75089;  
AC  
XX 28-JAN-1999 (first entry)  
DT  
XX Human secreted protein encoded by gene 33 clone HEMD84.  
DE  
XX Human; secreted protein; fusion protein; gene therapy; protein therapy;  
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;  
KW developmental abnormality; foetal deficiency; blood; allergy; renal;  
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;  
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;  
KW cognitive disorder; schizophrenia; prostate; osteoarthritis; thyroid; digestion;  
KW osteoporosis; arthritis; testis; lung; chylodiditis; thyroid; digestion;  
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.  
XX  
OS Homo sapiens.  
XX  
XX MO9839446-A2.  
XX  
XX 11-SEP-1998.  
XX  
XX 06-MAR-1998; 98MO-US04492.  
XX  
XX 07-MAR-1997; 97US-0038621.  
XX 07-MAR-1997; 97US-0040161.  
XX 07-MAR-1997; 97US-0040162.  
XX 07-MAR-1997; 97US-0040163.  
XX 07-MAR-1997; 97US-0040333.  
XX 07-MAR-1997; 97US-0040334.  
XX 07-MAR-1997; 97US-0040336.  
XX 07-MAR-1997; 97US-0040626.  
XX 11-APR-1997; 97US-0043311.  
XX 11-APR-1997; 97US-0043312.  
XX 11-APR-1997; 97US-0043313.  
XX 11-APR-1997; 97US-0043314.  
XX 11-APR-1997; 97US-0043315.  
XX 11-APR-1997; 97US-0043568.  
XX 11-APR-1997; 97US-0043569.  
XX 11-APR-1997; 97US-0043576.  
XX 11-APR-1997; 97US-0043578.  
XX 11-APR-1997; 97US-0043580.  
XX 11-APR-1997; 97US-0043668.  
XX 11-APR-1997; 97US-0043670.  
XX 11-APR-1997; 97US-0043671.  
XX 11-APR-1997; 97US-0043672.  
XX 11-APR-1997; 97US-0043674.  
XX 23-MAY-1997; 97US-0047492.  
XX 23-MAY-1997; 97US-0047500.  
XX 23-MAY-1997; 97US-0047501.  
XX 23-MAY-1997; 97US-0047502.  
XX 23-MAY-1997; 97US-0047503.  
XX 23-MAY-1997; 97US-0047581.  
XX 23-MAY-1997; 97US-0047582.  
XX 23-MAY-1997; 97US-0047583.  
XX 23-MAY-1997; 97US-0047584.

PR 23-MAY-1997; 97US-0047585.  
PR 23-MAY-1997; 97US-0047586.  
PR 23-MAY-1997; 97US-0047587.  
PR 23-MAY-1997; 97US-0047588.  
PR 23-MAY-1997; 97US-0047589.  
PR 23-MAY-1997; 97US-0047590.  
PR 23-MAY-1997; 97US-0047592.  
PR 23-MAY-1997; 97US-0047593.  
PR 23-MAY-1997; 97US-0047594.  
PR 23-MAY-1997; 97US-0047595.  
PR 23-MAY-1997; 97US-0047596.  
PR 23-MAY-1997; 97US-0047597.  
PR 23-MAY-1997; 97US-0047598.  
PR 23-MAY-1997; 97US-0047599.  
PR 23-MAY-1997; 97US-0047600.  
PR 23-MAY-1997; 97US-0047601.  
PR 23-MAY-1997; 97US-0047612.  
PR 23-MAY-1997; 97US-0047613.  
PR 23-MAY-1997; 97US-0047614.  
PR 23-MAY-1997; 97US-0047615.  
PR 23-MAY-1997; 97US-0047617.  
PR 23-MAY-1997; 97US-0047618.  
PR 23-MAY-1997; 97US-0047632.  
PR 23-MAY-1997; 97US-0047633.  
PR 06-JUN-1997; 97US-0048964.  
PR 06-JUN-1997; 97US-0048974.  
PR 22-AUG-1997; 97US-0056630.  
PR 22-AUG-1997; 97US-0056631.  
PR 22-AUG-1997; 97US-0056632.  
PR 22-AUG-1997; 97US-0056636.  
PR 22-AUG-1997; 97US-0056637.  
PR 22-AUG-1997; 97US-0056662.  
PR 22-AUG-1997; 97US-0056664.  
PR 22-AUG-1997; 97US-0056845.  
PR 22-AUG-1997; 97US-0056862.  
PR 22-AUG-1997; 97US-0056864.  
PR 22-AUG-1997; 97US-0056872.  
PR 22-AUG-1997; 97US-0056874.  
PR 22-AUG-1997; 97US-0056875.  
PR 22-AUG-1997; 97US-0056876.  
PR 22-AUG-1997; 97US-0056877.  
PR 22-AUG-1997; 97US-0056878.  
PR 22-AUG-1997; 97US-0056879.  
PR 22-AUG-1997; 97US-0056880.  
PR 22-AUG-1997; 97US-0056881.  
PR 22-AUG-1997; 97US-0056882.  
PR 22-AUG-1997; 97US-0056884.  
PR 22-AUG-1997; 97US-0056886.  
PR 22-AUG-1997; 97US-0056887.  
PR 22-AUG-1997; 97US-0056888.  
PR 22-AUG-1997; 97US-0056889.  
PR 22-AUG-1997; 97US-0056892.  
PR 22-AUG-1997; 97US-0056893.  
PR 22-AUG-1997; 97US-0056894.  
PR 22-AUG-1997; 97US-0056903.  
PR 22-AUG-1997; 97US-0056908.  
PR 22-AUG-1997; 97US-0056909.  
PR 22-AUG-1997; 97US-0056910.  
PR 22-AUG-1997; 97US-0056911.  
PR 05-SEP-1997; 97US-0057650.  
PR 05-SEP-1997; 97US-0057761.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Bednarik DP, Brewer LA, Carter KC, Duan R, Ebner R, Endress GA;  
PI Feng P, Ferris AM, Fischer CL, Graves KA, Greene JM, Hu JS, CA;  
PI Kyaw H, Lafleur DW, Li Y, Moore PA, Ni J, Olsen HS, Rosen CA;  
PI Ruben SM, Shi Y, Soppet DR, Young PE, Yu GL, Zeng Z;  
XX WPI: 1998-609887/51.  
XX N-PSDB; AAV34186.  
XX  
XX New isolated human genes and the secreted polypeptides they encode



CC The invention relates to human polynucleotides (AA179941-AA193841) and  
CC the encoded proteins (AA000010-AA03910) that exhibit activity elating to  
CC cytokine, cell proliferation or cell differentiation or which may induce

CC AAF15566 to AAF16505 encode the human prostate cancer associated  
CC proteins, called prostate cancer antigens, given in AAB55363 to AAB57302  
CC The prostate cancer antigens can have neuroprotective, cytoskeletal,  
CC cardioactive, immunomodulatory, muscular, vlnetrary, gastrointestinal,  
CC neurotropic, antiinfective, gynecological and antibacterial activities  
CC and can be used in gene therapy. The prostate cancer antigen  
CC polynucleotides may be used for detection of prostate cancer, chromosome  
CC identification, as chromosome markers, and for numerous other diagnostic  
CC or research purposes. The prostate cancer antigens may be used to treat

CC disorders such as neural, immune, muscular, reproductive,  
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative  
CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to  
CC AAF57303 represent sequences used in the exemplification of the present  
CC invention.

XX Sequence 146 AA;

Query Match 20.0%; Score 6; DB 21; Length 146;

Best Local Similarity 100.0%; Pred. No. 71;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 17 YISGQI 22  
|||||  
Db 41 YISGQI 46

## RESULT 26

ABP26232  
ID ABP26232 standard; Protein; 163 AA.

XX ABP26232;

DT 02-JUL-2002 (first entry)

DE Streptococcus polypeptide SEQ ID NO 1640.

XX Streptococcus; GAS; group B streptococcus; Streptococcus agalactiae;

XX group A streptococcus; Streptococcus pyogenes; antibacterial;  
XX antiinflammatory; infection; vaccine; meningitis; gene therapy.

OS Streptococcus pyogenes.

PN WO200234771-A2.

XX 02-MAY-2002.

PF 29-OCT-2001; 2001WO-GB04789.

XX 27-OCT-2000; 2000GB-0026333.

PR 24-NOV-2000; 2000GB-0028727.

PR 07-MAR-2001; 2001GB-0005640.

XX (CHIR-) CHIRON SPA.  
PA (GENO-) INST GENOMIC RES.

XX Telford J, Maignani V, Margalit Ros YI, Grandi G, Fraser C;  
PI Tectelin H;

XX WPI; 2002-352536/38.  
R N-PSDB; AAN66863.

XX New Streptococcus protein for the treatment or prevention of infection  
PT or disease caused by Streptococcus bacteria, such as meningitis, and  
PT for detecting a compound that binds to the protein -

PS Claim 1; Page 3315; 4525pp; English.

XX The invention relates to a protein (ABP25413-ABP30895) from group B  
CC streptococcus/GAS (Streptococcus agalactiae) or group A streptococcus/GAS  
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in  
CC the specification. The proteins have antibacterial and antiinflammatory  
CC activity. (I), nucleic acids encoding (I), AAN66044-ABN71526 and  
CC antibodies that bind (I) are used in the manufacture of medicaments for  
CC the treatment or prevention of infection or disease caused by  
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.  
CC Nucleic acids encoding (I) is used to detect Streptococcus in a  
CC biological sample. (I) is used to determine whether a compound binds to  
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be  
CC used as a vaccine or diagnostic composition. The disease caused by  
CC Streptococcus that is prevented or treated may be meningitis. Nucleic  
CC acid encoding (I) may be used to recombinantly produce (I) and may be  
CC used in gene therapy. Antibodies to (I) are used for affinity

CC chromatography, immunoassays, and distinguishing/identifying  
CC Streptococcus proteins.

XX Sequence 163 AA;

Query Match 20.0%; Score 6; DB 23; Length 163;

Best Local Similarity 100.0%; Pred. No. 79;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 SLPELD 8  
|||||  
Db 139 SLPELD 144

## RESULT 27

AAM53253  
ID AAM53253 standard; Protein; 188 AA.

XX AAM53253;

DT 30-JUL-1998 (first entry)

DE Candida albicans fungal antigen - allergen SEQ ID NO:6.

XX Candida albicans; vaccine; allergen; antigenic protein; fungal antigen;

XX immune response; infection; insoluble.

XX Candida albicans.

XX WO9809990-A1.

XX 12-MAR-1998.

XX 29-AUG-1997; 97WO-JP03041.

XX 31-MAR-1997; 97JP-0099775.

XX 04-SEP-1996; 96JP-0255400.

XX (TAKI) TAKARA SHUZO CO LTD.

XX Endo M, Kato I, Mizutani S, Takesako K;

XX WPI; 1998-193553/17.  
DR N-PSDB; AAV20822.

XX Fungal antigens comprising insoluble fraction of fungal cells -  
PT useful for, e.g. stimulating immune response and treatment and  
PT diagnosis of fungal infection(s)

XX Claim 28; Page 79-80; 108pp; Japanese.

XX The present sequence represents an active vaccine component or allergen  
CC derived from Candida albicans, which is an antigenic protein. The  
CC present invention describes fungal antigens, comprising the insoluble  
CC fraction of fungal cells having completely/partially removed cell walls.  
CC Also described are nucleic acids encoding the antigens and a method for  
CC producing the antigens. The antigens can be used for preparing  
CC therapeutic compositions for stimulating immune response, e.g. as a  
CC vaccine. They can be used for treatment of fungal infections, treatment  
CC and prevention of allergies and diagnosis of fungal infections in  
CC vertebrates. The vaccines are not live, and have low toxicity.

XX Sequence 188 AA;

Query Match 20.0%; Score 6; DB 19; Length 188;

Best Local Similarity 100.0%; Pred. No. 91;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 ATEPYI 18  
|||||  
Db 1 ATEPYI 6

RESULT 28  
 AAG92973 standard; Protein; 200 AA.  
 ID AAG92973 standard; Protein; 200 AA.  
 XX  
 AC AAG92973;  
 XX  
 DT 26-SEP-2001 (first entry)  
 XX  
 DE C glutamicum protein fragment SEQ ID NO: 6727.  
 XX  
 KM Corynebacterium; amino acid synthesis; vitamin; saccharide;  
 KM organic acid synthesis.  
 XX  
 OS Corynebacterium glutamicum.  
 XX  
 PN EP108790-A2.  
 XX  
 PD 20-JUN-2001.  
 XX  
 PR 18-DEC-2000; 2000EP-0127688.  
 XX  
 PR 16-DEC-1999; 99JP-0377484.  
 PR 07-APR-2000; 2000JP-0159162.  
 PR 03-AUG-2000; 2000JP-0280988.  
 XX  
 PA (KYOW ) KYOWA HAKKO KOGYO KK.  
 XX  
 PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;  
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;  
 DR WPI; 2001-376931/40.  
 DR N-PSDB; AAH68192.  
 PT Novel polynucleotides derived from Corynebacterium bacteria, for identifying  
 PT mutation point of a gene, measuring expression of a gene, analysing  
 PT expression profile or pattern of a gene and identifying homologous gene  
 PT  
 XX  
 PS Claim 17; SEQ ID NO: 6727; 246bp + Sequence Listing; English.  
 XX  
 CC The present invention provides a number of nucleotide and protein  
 CC sequences from the Corynebacterium bacterium Corynebacterium glutamicum. These  
 CC are useful for identifying the mutation point of a gene derived from a  
 CC mutant of corynebacterium bacterium, measuring expression amount and  
 CC analysing the expression profile or expression pattern of a gene derived  
 CC from Corynebacterium bacterium, and identifying a homologue of a gene derived  
 CC from corynebacterium bacterium. Corynebacterium bacteria are useful for producing  
 CC amino acids, nucleic acids, vitamins, saccharides and organic acids,  
 CC particularly L-lysine. The present sequence is a protein described  
 CC in the exemplification of the invention.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from the  
 CC European Patent Office.  
 CC  
 SQ Sequence 200 AA;  
 XX  
 Query Match 20.0%; Score 6; DB 22; Length 200;  
 Best Local Similarity 100.0%; Pred. No. 96;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 LPELDY 9  
 Db 6 LPELDY 11  
 XX  
 RESULT 29  
 ID AAU00514 standard; Protein; 200 AA.  
 XX  
 AC AAU00514;  
 XX  
 DT 09-MAY-2001 (first entry)  
 XX

DE C. melassecola superoxide dismutase.  
 XX  
 KM Superoxide dismutase; sod; corynebacterium microorganism; metabolite;  
 KM vitamin; D-pantothenic acid; L-lysine; amplification; animal nutrition.  
 XX  
 OS Corynebacterium melassecola.  
 XX  
 PN EP1077261-A2.  
 XX  
 PD 21-FEB-2001.  
 XX  
 PR 02-AUG-2000; 2000EP-0116669.  
 XX  
 PR 13-AUG-1999; 99US-0373731.  
 XX  
 PA (DEGS ) DEGUSSA-HUELS AG.  
 PA (CNRS ) CNRS CENT NAT RECH SCI.  
 XX  
 PI Merkmann M, Guyonvarch A, Marx A;  
 XX  
 DR WPI; 2001-212718/22.  
 DR N-PSDB; AAS00520.  
 XX  
 PT New DNA encoding superoxide dismutase of Corynebacterium, useful for  
 PT producing transformants with increased production of metabolites,  
 PT particularly lysine -  
 XX  
 PS Claim 3; Page 14-15; 19pp; English.  
 XX  
 CC The sequence represents Corynebacterium melassecola superoxide dismutase  
 CC (sod) polypeptide. Corynebacterium microorganisms may be transformed with a  
 CC sod DNA sequence and the sod gene can be amplified. Sod is then often  
 CC overexpressed in corynebacterium bacteria. Corynebacterium that overexpress sod  
 CC are used for production of metabolites, particularly nucleotides,  
 CC vitamins and amino acids, especially D-pantothenic acid or, specifically,  
 CC L-lysine. These metabolites are useful in human or animal nutrition and  
 CC as pharmaceuticals. Overexpression of sod in corynebacterium increases the  
 CC yield of particular metabolites.  
 CC  
 SQ Sequence 200 AA;  
 XX  
 Query Match 20.0%; Score 6; DB 22; Length 200;  
 Best Local Similarity 100.0%; Pred. No. 96;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 LPELDY 9  
 Db 6 LPELDY 11  
 XX  
 RESULT 30  
 ID ABB47790 standard; Protein; 207 AA.  
 XX  
 AC ABB47790;  
 XX  
 DT 05-FEB-2002 (first entry)  
 XX  
 DE Listeria monocytogenes protein #494.  
 XX  
 KM Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;  
 KM vitamin B12; bacterial infection; disease.  
 XX  
 OS Listeria monocytogenes.  
 XX  
 PN WO200177335-A2.  
 XX  
 PD 18-OCT-2001.  
 XX  
 PR 11-APR-2001; 2001WO-FR01118.  
 XX  
 PR 11-APR-2000; 2000FR-0004629.  
 XX

PA (INSP ) INST PASTEUR.

XX Buchrieser C, Frangeul L, Couve E, Rusniok C, Feihl H, Dehoux P,  
 PI Duesuguet O, Chetouani F, Nedjari H, Glaeser P, Kunst R, Cossart P,  
 PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA,  
 PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A,  
 PI Chakraborty T, Domann F, Hain T, Berche P, Charbit A, Durant L,  
 PI Perez-Diaz J, Baguero F, Garcia Del Portillo F, Gomez-Lopez N,  
 PI Madueno E, De Pablos B, Wehlend J, Kaerst U, Entian K, Hauf J,  
 PI Rose M, Voss H;  
 DR WPI: 2002-010914/01.

XX Genomic sequence for *Listeria monocytogenes*, useful e.g. for treatment  
 PT and prevention of *Listeria* and related bacterial infections, and  
 PT related polypeptides -

XX Claim 6; SEQ ID No 495; 192pp; French.

XX The present invention relates to the genome sequence of *Listeria*  
 CC *monocytogenes* EGD-e (see ABA03041). The genome sequence and fragments of  
 CC it are useful for selecting probes and primers for detecting genes in *L.*  
 CC *monocytogenes* and related organisms, and for studying genetic  
 CC polymorphisms and other genomes. The present sequence is a protein  
 CC encoded by the genome sequence of the present invention. Proteins  
 CC expressed from the genome sequence are useful for raising specific  
 CC antibodies, identification of *L. monocytogenes* and related organisms, and  
 CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin  
 CC B12. The genome sequence and proteins encoded by it are also useful for  
 CC selecting compounds that regulate gene expression and cell replication  
 CC and modulate *L. monocytogenes*-related diseases. In addition, the genome  
 CC sequence and proteins encoded by it are useful in pharmaceutical and  
 CC vaccines compositions for the treatment or prevention of infections by *L.*  
 CC *monocytogenes* and related organisms.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIP0  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 207 AA;

XX Query Match 20.0%; Score 6; DB 23; Length 207;  
 XX Best Local Similarity 100.0%; Pred. No. 99;  
 XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 PYISGO 21  
 |||||  
 Db 191 PYISGO 196

XX RESULT 31  
 XX AAY34657  
 ID AAY34657 standard; Protein; 214 AA.

XX AAY34657;  
 XX 13-SEP-1999 (first entry)

XX C. pneumoniae protein involved in intermediate metabolism.

XX Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;  
 KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;  
 KW vaccine; neutralising epitope.

XX Chlamydia pneumoniae.

XX W09927105-A2.

XX 03-JUN-1999.

XX 20-NOV-1998; 98WO-1B01890.

XX 04-NOV-1998; 98US-0107078.  
 PR 21-NOV-1997; 97FR-0014673.

XX (GEST ) GENSET.

XX Griffiths R;

XX WPI: 1999-357842/30.

XX Genome sequence of *Chlamydia pneumoniae*  
 PT Page 665; Disclosure; 1912pp; English.

XX AAY34584-Y35879 represent the proteins encoded by all the open reading  
 CC frames in the complete genome (see AAY31990) of *Chlamydia pneumoniae*.  
 CC C. pneumoniae causes respiratory diseases such as pneumonia and  
 CC bronchitis and is thought to be a contributing factor in heart  
 CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema  
 CC nodosum or pharyngitis. The polypeptides encoded by the open reading  
 CC frames of the C. pneumoniae genome (see AAY34584-Y35879) can be used in  
 CC immunogenic compositions as vaccines. Vectors containing C. pneumoniae  
 CC nucleotide sequences can also be used as immunogenic compositions,  
 CC especially where the vector directs the expression of a neutralising  
 CC epitope of C. pneumoniae.

XX Sequence 214 AA;

XX Query Match 20.0%; Score 6; DB 20; Length 214;  
 XX Best Local Similarity 100.0%; Pred. No. 1e+02;  
 XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 YSLPEL 7  
 |||||  
 Db 13 YSLPEL 18

XX RESULT 32

XX AAB94056  
 ID AAB94056 standard; Protein; 221 AA.

XX AAB94056;

XX 26-JUN-2001 (first entry)

XX Human protein sequence SEQ ID NO:14228.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

XX 27-AUG-1999; 99JP-0300253.

XX 11-JAN-2000; 2000JP-0118776.

XX 02-MAY-2000; 2000JP-0183767.

XX 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Negai K, Otsuki T;

XX WPI: 2001-318749/34.  
 XX Primer sets for synthesizing polynucleotides, particularly the 5602  
 PT full-length cDNAs defined in the specification, and for the detection  
 PT and/or diagnosis of the abnormality of the proteins encoded by the  
 PT full-length cDNAs -

PS Claim 8; SEQ ID 14228; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602  
CC full-length cDNAs defined in the specification. Where a primer set  
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
CC to the complementary strand of a polynucleotide which comprises one of  
CC the 5602 nucleotide sequences defined in the specification, where the  
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
CC of an oligonucleotide comprising a sequence complementary to the  
CC complementary strand of a polynucleotide which comprises a 5'-end  
CC sequence and an oligonucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in  
CC the specification. The primer sets can be used in antisense therapy and  
CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH03166 to AAH1628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
CC represent oligonucleotides, all of which are used in the exemplification  
CC of the present invention.

XX  
SQ Sequence 221 AA;  
Query Match 20.0%; Score 6; DB 22; Length 221;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 ATEPYI 18  
|||  
Db 211 ATEPYI 216

RESULT 33  
AAG78026  
ID AAG78026 standard; Protein; 226 AA.  
XX  
AC AAG78026;  
XX  
DT 15-JAN-2002 (first entry)  
XX  
DE Piscirickettsia salmonis ICME.  
XX  
KW Piscirickettsia salmonis; piscirickettsiosis; salmonid rickettsial;  
KW septiccaemia; SRS; surface antigen; vaccine; antibacterial; fish;  
XX type strain LF-89.  
XX  
OS Piscirickettsia salmonis.  
XX  
PN WO200168865-A2.  
XX  
PD 20-SEP-2001.  
XX  
PF 12-MAR-2001; 2001WO-GB01055.  
XX  
PR 11-MAR-2000; 2000GB-0005838.  
PR 01-JUL-2000; 2000GB-0016080.  
PR 01-JUL-2000; 2000GB-0016082.  
PR 29-JUL-2000; 2000GB-0018599.  
XX  
PA (AQUA-) AQUA HEALTH EURO LTD.  
XX  
PI Simard N, Brouwers H, Jones S, Griffiths S, Valenzuela P,  
PI Burzio L;  
XX  
XX WPI; 2001-639050/73.  
DR N-PSDB; AAH79041.  
XX  
XX New nucleic acids encoding an amino acid sequence homologous to the  
PT surface antigen present on Piscirickettsia salmonis are useful to  
PT protect fish against piscirickettsiosis -

XX  
PS Claim 6; Fig 6; 25pp; English.  
XX  
XX The invention relates to nucleic acid sequences and the encoded protein  
CC of a least part of the surface antigen present on Piscirickettsia  
CC salmonis for production of a vaccine with antibacterial activity to  
CC protect fish against P. salmonis which causes piscirickettsiosis, also  
CC known as salmonid rickettsial septiccaemia.  
XX  
SQ Sequence 226 AA;  
Query Match 20.0%; Score 6; DB 22; Length 226;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SLPEID 8  
|||  
Db 98 SLPEID 103

RESULT 34  
AAG49754  
ID AAG49754 standard; Protein; 294 AA.  
XX  
AC AAG49754;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 62974.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0135621.  
PR 27-MAY-1999; 99US-0135632.  
PR 28-MAY-1999; 99US-01356782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0138119.  
PR 16-JUN-1999; 99US-0138452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 16-AUG-1999; 99US-0148684.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160757.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160770.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.

PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.  
Query Match 20.0%; Score 6; DB 21; Length 294;  
Best Local Similarity 100.0%; Pred. No.1.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 13 ATEPYI 18  
Db 20 ATEPYI 25  
RESULT 35  
AAG50274  
ID AAG50274 standard; Protein; 294 AA.  
XX AAG50274;  
AC 18-OCT-2000 (first entry)  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 63691.  
DE Protein identification; signal transduction pathway; metabolic pathway;  
XX hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.  
XX Arabidopsis thaliana.  
OS EPI033405-A2.  
XX 06-SEP-2000.  
XX 25-FEB-2000; 2000EP-0301439.  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139464.  
PR 18-JUN-1999; 99US-0139465.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 28-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147312.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.

PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148177.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158368.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 20.0% Score 6, DB 21, Length 294;  
Best Local Similarity 100.0%; Pred.No.1,4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 13 ATEPYI 18  
Db 20 ATEPYI 25  
RESULT 36  
AAG29351  
ID AAG29351 standard; Protein; 311 AA.  
XX  
AC AAG29351;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 34908.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hydrolisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.  
OS Arabidopsis thaliana.  
XX  
PN EPI033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PE 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 18-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.



PR 17-JUN-1999; 99US-0139492.  
 PR 18-JUN-1999; 99US-0139454.  
 PR 18-JUN-1999; 99US-0139455.  
 PR 18-JUN-1999; 99US-0139456.  
 PR 18-JUN-1999; 99US-0139457.  
 PR 18-JUN-1999; 99US-0139458.  
 PR 18-JUN-1999; 99US-0139459.  
 PR 18-JUN-1999; 99US-0139460.  
 PR 18-JUN-1999; 99US-0139461.  
 PR 18-JUN-1999; 99US-0139462.  
 PR 18-JUN-1999; 99US-0139463.  
 PR 18-JUN-1999; 99US-0139750.  
 PR 18-JUN-1999; 99US-0139753.  
 PR 21-JUN-1999; 99US-0139817.  
 PR 22-JUN-1999; 99US-0139899.  
 PR 23-JUN-1999; 99US-0140353.  
 PR 23-JUN-1999; 99US-0140354.  
 PR 24-JUN-1999; 99US-0140695.  
 PR 28-JUN-1999; 99US-0140823.  
 PR 29-JUN-1999; 99US-0140991.  
 PR 30-JUN-1999; 99US-0141287.  
 PR 01-JUL-1999; 99US-0141842.  
 PR 01-JUL-1999; 99US-0141842.  
 PR 02-JUL-1999; 99US-0142154.  
 PR 06-JUL-1999; 99US-0142300.  
 PR 08-JUL-1999; 99US-0142803.  
 PR 09-JUL-1999; 99US-0142920.  
 PR 12-JUL-1999; 99US-0142977.  
 PR 13-JUL-1999; 99US-0143542.  
 PR 14-JUL-1999; 99US-0143624.  
 PR 15-JUL-1999; 99US-0144005.  
 PR 16-JUL-1999; 99US-0144085.  
 PR 16-JUL-1999; 99US-0144086.  
 PR 19-JUL-1999; 99US-0144325.  
 PR 19-JUL-1999; 99US-0144331.  
 PR 19-JUL-1999; 99US-0144332.  
 PR 19-JUL-1999; 99US-0144333.  
 PR 19-JUL-1999; 99US-0144334.  
 PR 20-JUL-1999; 99US-0144335.  
 PR 20-JUL-1999; 99US-0144352.  
 PR 20-JUL-1999; 99US-0144632.  
 PR 20-JUL-1999; 99US-0144884.  
 PR 21-JUL-1999; 99US-0144814.  
 PR 21-JUL-1999; 99US-0145086.  
 PR 21-JUL-1999; 99US-0145088.  
 PR 22-JUL-1999; 99US-0145085.  
 PR 22-JUL-1999; 99US-0145087.  
 PR 22-JUL-1999; 99US-0145089.  
 PR 22-JUL-1999; 99US-0145192.  
 PR 23-JUL-1999; 99US-0145218.  
 PR 23-JUL-1999; 99US-0145218.  
 PR 23-JUL-1999; 99US-0145224.  
 PR 26-JUL-1999; 99US-0145276.  
 PR 27-JUL-1999; 99US-0145913.  
 PR 27-JUL-1999; 99US-0145918.  
 PR 28-JUL-1999; 99US-0145951.  
 PR 02-AUG-1999; 99US-0146386.  
 PR 02-AUG-1999; 99US-0146388.  
 PR 02-AUG-1999; 99US-0146388.  
 PR 03-AUG-1999; 99US-0146389.  
 PR 03-AUG-1999; 99US-0147038.  
 PR 04-AUG-1999; 99US-0147204.  
 PR 04-AUG-1999; 99US-0147302.  
 PR 05-AUG-1999; 99US-0147192.  
 PR 05-AUG-1999; 99US-0147260.  
 PR 06-AUG-1999; 99US-0147303.  
 PR 06-AUG-1999; 99US-0147416.  
 PR 09-AUG-1999; 99US-0147493.  
 PR 09-AUG-1999; 99US-0147935.  
 PR 10-AUG-1999; 99US-0148171.  
 PR 11-AUG-1999; 99US-0148319.  
 PR 12-AUG-1999; 99US-0148341.  
 PR 13-AUG-1999; 99US-0148565.

PR 13-AUG-1999; 99US-0148684.  
 PR 16-AUG-1999; 99US-0149368.  
 PR 17-AUG-1999; 99US-0149175.  
 PR 18-AUG-1999; 99US-0149426.  
 PR 20-AUG-1999; 99US-0149722.  
 PR 20-AUG-1999; 99US-0149723.  
 PR 20-AUG-1999; 99US-0149729.  
 PR 23-AUG-1999; 99US-0149902.  
 PR 23-AUG-1999; 99US-0149902.  
 PR 25-AUG-1999; 99US-0150566.  
 PR 26-AUG-1999; 99US-0150884.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 27-AUG-1999; 99US-0151066.  
 PR 27-AUG-1999; 99US-0151080.  
 PR 30-AUG-1999; 99US-0151303.  
 PR 31-AUG-1999; 99US-0151438.  
 PR 01-SEP-1999; 99US-0151930.  
 PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 22-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 23-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 20.0%; Score 6; DB 21; Length 311;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 13 ATEPYI 18  
 |||||  
 Db 38 ATEPYI 43

RESULT 37

AA649753  
ID AA649753 standard; Protein; 312 AA.  
XX  
AC AA649753;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 62973.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridization assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
PN EPI033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
25-FEB-1999; 99US-0121825.  
XX 05-MAR-1999; 99US-0121180.  
XX 09-MAR-1999; 99US-0123548.  
XX 23-MAR-1999; 99US-0125788.  
XX 25-MAR-1999; 99US-0126264.  
XX 29-MAR-1999; 99US-0126785.  
XX 01-APR-1999; 99US-0126462.  
XX 06-APR-1999; 99US-0128234.  
XX 08-APR-1999; 99US-0128714.  
XX 16-APR-1999; 99US-0129845.  
XX 19-APR-1999; 99US-0130077.  
XX 21-APR-1999; 99US-0130449.  
XX 23-APR-1999; 99US-0130510.  
XX 23-APR-1999; 99US-0130891.  
XX 28-APR-1999; 99US-0131449.  
XX 30-APR-1999; 99US-0132048.  
XX 30-APR-1999; 99US-0132407.  
XX 04-MAY-1999; 99US-0132484.  
XX 05-MAY-1999; 99US-0132485.  
XX 06-MAY-1999; 99US-0132486.  
XX 06-MAY-1999; 99US-0132487.  
XX 07-MAY-1999; 99US-0132863.  
XX 11-MAY-1999; 99US-0134256.  
XX 14-MAY-1999; 99US-0134218.  
XX 14-MAY-1999; 99US-0134219.  
XX 14-MAY-1999; 99US-0134221.  
XX 14-MAY-1999; 99US-0134370.  
XX 18-MAY-1999; 99US-0134768.  
XX 19-MAY-1999; 99US-0134941.  
XX 20-MAY-1999; 99US-0135124.  
XX 21-MAY-1999; 99US-0135353.  
XX 24-MAY-1999; 99US-0135629.  
XX 25-MAY-1999; 99US-0136021.  
XX 27-MAY-1999; 99US-0136392.  
XX 28-MAY-1999; 99US-0136782.  
XX 01-JUN-1999; 99US-0137222.  
XX 03-JUN-1999; 99US-0137528.  
XX 04-JUN-1999; 99US-0137502.  
XX 07-JUN-1999; 99US-0137724.  
XX 08-JUN-1999; 99US-0138094.  
XX 10-JUN-1999; 99US-0138540.  
XX 10-JUN-1999; 99US-0138647.  
XX 14-JUN-1999; 99US-0139110.  
XX 16-JUN-1999; 99US-0139452.  
XX 16-JUN-1999; 99US-0139453.  
XX 17-JUN-1999; 99US-0139492.  
XX 18-JUN-1999; 99US-0139454.  
XX 18-JUN-1999; 99US-0139455.  
XX 18-JUN-1999; 99US-0139456.  
XX 18-JUN-1999; 99US-0139457.  
XX 18-JUN-1999; 99US-0139458.  
XX 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140585.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142380.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 22-JUL-1999; 99US-0145193.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.

PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154039.  
PR 16-SEP-1999; 99US-0154779.  
PR 20-SEP-1999; 99US-0155139.  
PR 22-SEP-1999; 99US-0155486.  
PR 23-SEP-1999; 99US-0155659.  
PR 24-SEP-1999; 99US-0156458.  
PR 28-SEP-1999; 99US-0156596.  
PR 29-SEP-1999; 99US-0157117.  
PR 04-OCT-1999; 99US-0157753.  
PR 05-OCT-1999; 99US-0157865.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159688.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160981.  
PR 21-OCT-1999; 99US-0160989.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 26-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 28-OCT-1999; 99US-0162142.  
PR 29-OCT-1999; 99US-0162142.

Query Match 20.0%; Score 6; DB 21; Length 312;  
Best Local Similarity 100.0%; Pred. No.1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 ATEPYI 18  
|||  
Db 38 ATEPYI 43

RESULT 38  
AAG50273  
ID AAG50273 standard; Protein; 312 AA.

XX AAG50273;  
AC  
DT 18-OCT-2000 (first entry)  
XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 63690.  
XX  
XX Protein identification; signal transduction pathway; metabolic pathway;  
XX hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.  
OS Arabidopsis thaliana.  
XX  
XX EPI033405-A2.  
XX  
XX PD 06-SEP-2000.  
XX  
XX PF 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0128845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132484.  
PR 06-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134231.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899;  
PR 23-JUN-1999; 99US-0140353;  
PR 23-JUN-1999; 99US-0140354;  
PR 24-JUN-1999; 99US-0140695;  
PR 28-JUN-1999; 99US-0140823;  
PR 29-JUN-1999; 99US-0140991;  
PR 30-JUN-1999; 99US-0141287;  
PR 01-JUL-1999; 99US-0141842;  
PR 01-JUL-1999; 99US-0142154;  
PR 02-JUL-1999; 99US-0142055;  
PR 06-JUL-1999; 99US-0142390;  
PR 08-JUL-1999; 99US-0142803;  
PR 09-JUL-1999; 99US-0142920;  
PR 12-JUL-1999; 99US-0142977;  
PR 13-JUL-1999; 99US-0143542;  
PR 14-JUL-1999; 99US-0143624;  
PR 15-JUL-1999; 99US-0144005;  
PR 16-JUL-1999; 99US-0144085;  
PR 16-JUL-1999; 99US-0144086;  
PR 19-JUL-1999; 99US-0144325;  
PR 19-JUL-1999; 99US-0144331;  
PR 19-JUL-1999; 99US-0144332;  
PR 19-JUL-1999; 99US-0144333;  
PR 19-JUL-1999; 99US-0144334;  
PR 19-JUL-1999; 99US-0144335;  
PR 20-JUL-1999; 99US-0144352;  
PR 20-JUL-1999; 99US-0144632;  
PR 20-JUL-1999; 99US-0144884;  
PR 21-JUL-1999; 99US-0144814;  
PR 21-JUL-1999; 99US-0145086;  
PR 21-JUL-1999; 99US-0145088;  
PR 22-JUL-1999; 99US-0145085;  
PR 22-JUL-1999; 99US-0145087;  
PR 22-JUL-1999; 99US-0145089;  
PR 22-JUL-1999; 99US-0145192;  
PR 23-JUL-1999; 99US-0145145;  
PR 23-JUL-1999; 99US-0145218;  
PR 26-JUL-1999; 99US-0145224;  
PR 26-JUL-1999; 99US-0145276;  
PR 27-JUL-1999; 99US-0145913;  
PR 27-JUL-1999; 99US-0145918;  
PR 27-JUL-1999; 99US-0145919;  
PR 28-JUL-1999; 99US-0145951;  
PR 02-AUG-1999; 99US-0146386;  
PR 02-AUG-1999; 99US-0146388;  
PR 02-AUG-1999; 99US-0146389;  
PR 03-AUG-1999; 99US-0147038;  
PR 04-AUG-1999; 99US-0147204;  
PR 04-AUG-1999; 99US-0147302;  
PR 05-AUG-1999; 99US-0147192;  
PR 05-AUG-1999; 99US-0147260;  
PR 06-AUG-1999; 99US-0147303;  
PR 06-AUG-1999; 99US-0147416;  
PR 09-AUG-1999; 99US-0147493;  
PR 09-AUG-1999; 99US-0147935;  
PR 10-AUG-1999; 99US-0148171;  
PR 11-AUG-1999; 99US-0148319;  
PR 12-AUG-1999; 99US-0148341;  
PR 13-AUG-1999; 99US-0148565;  
PR 13-AUG-1999; 99US-0148684;  
PR 16-AUG-1999; 99US-0149368;  
PR 17-AUG-1999; 99US-0149175;  
PR 18-AUG-1999; 99US-0149426;  
PR 20-AUG-1999; 99US-0149722;  
PR 20-AUG-1999; 99US-0149723;  
PR 20-AUG-1999; 99US-0149929;  
PR 23-AUG-1999; 99US-0149902;  
PR 23-AUG-1999; 99US-0149930;  
PR 25-AUG-1999; 99US-0150566;  
PR 26-AUG-1999; 99US-0150884;  
PR 27-AUG-1999; 99US-0151065;  
PR 27-AUG-1999; 99US-0151066;  
PR 27-AUG-1999; 99US-0151080;

PR 30-AUG-1999; 99US-0151303;  
PR 31-AUG-1999; 99US-0151438;  
PR 01-SEP-1999; 99US-0151930;  
PR 07-SEP-1999; 99US-0152363;  
PR 10-SEP-1999; 99US-0153070;  
PR 13-SEP-1999; 99US-0153758;  
PR 15-SEP-1999; 99US-0154018;  
PR 16-SEP-1999; 99US-0154039;  
PR 20-SEP-1999; 99US-0154779;  
PR 22-SEP-1999; 99US-0155139;  
PR 23-SEP-1999; 99US-0155486;  
PR 24-SEP-1999; 99US-0155649;  
PR 28-SEP-1999; 99US-0156458;  
PR 29-SEP-1999; 99US-0156596;  
PR 04-OCT-1999; 99US-0157117;  
PR 05-OCT-1999; 99US-0157753;  
PR 06-OCT-1999; 99US-0157865;  
PR 07-OCT-1999; 99US-0158029;  
PR 08-OCT-1999; 99US-0158232;  
PR 12-OCT-1999; 99US-0158369;  
PR 13-OCT-1999; 99US-0159293;  
PR 13-OCT-1999; 99US-0159294;  
PR 13-OCT-1999; 99US-0159295;  
PR 14-OCT-1999; 99US-0159329;  
PR 14-OCT-1999; 99US-0159330;  
PR 14-OCT-1999; 99US-0159331;  
PR 14-OCT-1999; 99US-0159637;  
PR 14-OCT-1999; 99US-0159638;  
PR 18-OCT-1999; 99US-0159584;  
PR 21-OCT-1999; 99US-0160741;  
PR 21-OCT-1999; 99US-0160767;  
PR 21-OCT-1999; 99US-0160768;  
PR 21-OCT-1999; 99US-0160770;  
PR 21-OCT-1999; 99US-0160814;  
PR 21-OCT-1999; 99US-0160815;  
PR 22-OCT-1999; 99US-0160960;  
PR 22-OCT-1999; 99US-0160981;  
PR 22-OCT-1999; 99US-0160989;  
PR 25-OCT-1999; 99US-0161404;  
PR 25-OCT-1999; 99US-0161405;  
PR 25-OCT-1999; 99US-0161406;  
PR 26-OCT-1999; 99US-0161359;  
PR 26-OCT-1999; 99US-0161360;  
PR 26-OCT-1999; 99US-0161361;  
PR 28-OCT-1999; 99US-0161920;  
PR 28-OCT-1999; 99US-0161992;  
PR 28-OCT-1999; 99US-0161993;  
PR 29-OCT-1999; 99US-0162142;

Query Match 20.0%; Score 6; DB 21; Length 312;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 ATEPYI 18  
|||||  
Db 38 ATEPYI 43

RESULT 39  
ABBS4867  
ID ABBS4867 standard; Protein; 338 AA.  
XX  
AC ABBS4867;  
XX  
DT 16-MAY-2002 (first entry)  
XX  
XX Lactococcus lactis protein purm.  
DE  
XX Biosynthesis, biodegradation; lactic bacterium; yogurt; cheese.  
KW  
XX Lactococcus lactis IL1403.  
OS  
XX  
PN FR2807446-A1.

XX 12-OCT-2001.  
 PD 11-APR-2000; 2000FR-0004630.  
 PF 11-APR-2000; 2000FR-0004630.  
 XX 11-APR-2000; 2000FR-0004630.  
 PR 11-APR-2000; 2000FR-0004630.  
 XX (INRG ) INRA INST NAT RECH AGRONOMICQUE.  
 PA Bolotine A, Sorokine A, Renault P, Ehrlich SD;  
 XX WPI; 2002-043418/06.  
 DR New nucleotide sequence useful in the identification or Lactococcus  
 PT lactis and related species -  
 XX  
 PS Claim 6; SEQ ID No 1569; 2504bp; French.  
 XX The present invention is related to a Lactococcus lactis nucleotide  
 CC sequence (AB90521) and related proteins (AB853300-AB855621). The  
 CC nucleic acid sequence is useful in the detection and/or amplification of  
 CC nucleic acid sequence, particularly to identify Lactococcus lactis or  
 CC related species. The proteins of the invention are useful for the  
 CC biosynthesis or biodegradation of a composition of interest. The  
 CC invention helps research in lactic bacteria, particularly useful in the  
 CC production of yogurt and cheese.  
 CC Note: The sequence data for this patent is based on equivalent patent  
 CC WO00017734 (published 18-OCT-2001) which is available in electronic  
 CC format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 338 AA;  
 Query Match 20.0%; Score 6; DB 23; Length 338;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 SLEPMD 8  
 Db 208 SLEPMD 213  
 RESULT 40  
 AAG29350  
 ID AAG29350 standard; Protein; 355 AA.  
 XX  
 AC AAG29350;  
 XX  
 XX 17-OCT-2000 (first entry)  
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 34907.  
 XX  
 KM Protein identification; signal transduction pathway; metabolic pathway;  
 KM hybridisation assay; genetic mapping; gene expression control; promoter;  
 KM termination sequence.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN EPI033405-A2.  
 XX  
 PD 06-SEP-2000.  
 PD  
 XX  
 PF 25-FEB-2000; 2000EP-0301439.  
 PF  
 XX 25-FEB-1999; 99US-0121825.  
 PR 05-MAR-1999; 99US-0123180.  
 PR 09-MAR-1999; 99US-0123548.  
 PR 23-MAR-1999; 99US-0125788.  
 PR 25-MAR-1999; 99US-0126264.  
 PR 29-MAR-1999; 99US-0126785.  
 PR 01-APR-1999; 99US-0127462.  
 PR 06-APR-1999; 99US-0128234.  
 PR 08-APR-1999; 99US-0128714.  
 PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.  
 PR 21-APR-1999; 99US-0130449.  
 PR 23-APR-1999; 99US-0130510.  
 PR 28-APR-1999; 99US-0130891.  
 PR 30-APR-1999; 99US-0131449.  
 PR 30-APR-1999; 99US-0132048.  
 PR 30-APR-1999; 99US-0132407.  
 PR 04-MAY-1999; 99US-0132484.  
 PR 05-MAY-1999; 99US-0132485.  
 PR 06-MAY-1999; 99US-0132486.  
 PR 06-MAY-1999; 99US-0132487.  
 PR 07-MAY-1999; 99US-0132853.  
 PR 11-MAY-1999; 99US-0134256.  
 PR 14-MAY-1999; 99US-0134218.  
 PR 14-MAY-1999; 99US-0134219.  
 PR 14-MAY-1999; 99US-0134221.  
 PR 14-MAY-1999; 99US-0134370.  
 PR 18-MAY-1999; 99US-0134768.  
 PR 19-MAY-1999; 99US-0134941.  
 PR 20-MAY-1999; 99US-0135124.  
 PR 21-MAY-1999; 99US-0135124.  
 PR 24-MAY-1999; 99US-0135353.  
 PR 25-MAY-1999; 99US-0135629.  
 PR 27-MAY-1999; 99US-0136021.  
 PR 28-MAY-1999; 99US-0136382.  
 PR 01-JUN-1999; 99US-0137222.  
 PR 03-JUN-1999; 99US-0137528.  
 PR 04-JUN-1999; 99US-0137502.  
 PR 07-JUN-1999; 99US-0137724.  
 PR 08-JUN-1999; 99US-0138094.  
 PR 10-JUN-1999; 99US-0138540.  
 PR 10-JUN-1999; 99US-0138847.  
 PR 14-JUN-1999; 99US-0139119.  
 PR 16-JUN-1999; 99US-0139452.  
 PR 16-JUN-1999; 99US-0139453.  
 PR 17-JUN-1999; 99US-0139452.  
 PR 18-JUN-1999; 99US-0139454.  
 PR 18-JUN-1999; 99US-0139455.  
 PR 18-JUN-1999; 99US-0139456.  
 PR 18-JUN-1999; 99US-0139457.  
 PR 18-JUN-1999; 99US-0139458.  
 PR 18-JUN-1999; 99US-0139459.  
 PR 18-JUN-1999; 99US-0139460.  
 PR 18-JUN-1999; 99US-0139461.  
 PR 18-JUN-1999; 99US-0139462.  
 PR 18-JUN-1999; 99US-0139463.  
 PR 18-JUN-1999; 99US-0139750.  
 PR 18-JUN-1999; 99US-0139763.  
 PR 21-JUN-1999; 99US-0139817.  
 PR 22-JUN-1999; 99US-0139899.  
 PR 23-JUN-1999; 99US-0140353.  
 PR 23-JUN-1999; 99US-0140354.  
 PR 24-JUN-1999; 99US-0140625.  
 PR 28-JUN-1999; 99US-0140823.  
 PR 29-JUN-1999; 99US-0140991.  
 PR 30-JUN-1999; 99US-0141287.  
 PR 01-JUL-1999; 99US-0141842.  
 PR 01-JUL-1999; 99US-0142154.  
 PR 02-JUL-1999; 99US-0142055.  
 PR 06-JUL-1999; 99US-0142390.  
 PR 08-JUL-1999; 99US-0142803.  
 PR 09-JUL-1999; 99US-0142920.  
 PR 12-JUL-1999; 99US-0142977.  
 PR 13-JUL-1999; 99US-0143542.  
 PR 14-JUL-1999; 99US-0143624.  
 PR 15-JUL-1999; 99US-0144005.  
 PR 16-JUL-1999; 99US-0144086.  
 PR 16-JUL-1999; 99US-0144086.  
 PR 19-JUL-1999; 99US-0144325.  
 PR 19-JUL-1999; 99US-0144331.  
 PR 19-JUL-1999; 99US-0144332.  
 PR 19-JUL-1999; 99US-0144333.  
 PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144864.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.

PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 20.0%; Score 6; DB 21; Length 355;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 ATEPYI 18  
|||||  
Db 82 ATEPYI 87

Search completed: April 9, 2003, 14:14:13  
Job time : 35.3119 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 8, 2003, 01:31:08 ; Search time 443.807 Seconds  
(without alignments)  
1967.259 Million cell updates/sec

Title: US-09-987-190-2  
Perfect score: 145  
Sequence: 1 KXSLPELDYFSAPEYISQINEIXYXX 30

Scoring table: BIOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Rgapop 6.0 , Rgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame+ p2n.model -DEV=xlh  
-O=/cgm2.1/USPTO.spool/US09987190/runat.02042003.092632.19221/app.query.fasta.1.526  
-DB=GenEmbl -OPMT=fastap -SUFFIX=rge -MITMATCH=0.1 -LOFCL=0 -LOFEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blisum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=40 -MODE=LOCAL  
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09987190 @CGN 1.1 1758 -runat.02042003.092632.19221 -NCPU=6 -ICPU=3  
-NO\_XLPEY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -RGAPOP=6 -RGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_hlg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pac.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sbs.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vt.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_om.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pac.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_sbs.\*  
28: em\_un.\*

29: em\_vi.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pln.\*  
35: em\_htg\_rtd.\*  
36: em\_htg\_man.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_htg\_hum.\*  
40: em\_htg\_mus.\*  
41: em\_hcgo\_other.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	141	97.2	1859	8 AF031478	AF031478 Candida a
2	107	73.8	2332	8 CSMSOD	Y11598 Candida sp.
3	103	73.5	735	1 NAU02341	U02341 Nocardi
4	101	69.7	666	1 AF333434	AF333434 Mycobacte
5	101	69.7	960	1 MAU11550	U11550 Mycobacteri
6	101	69.7	1741	1 AF180816	AF180816 Mycobacte
7	100	69.0	815	1 D13288	D13288 Mycobacteri
8	100	69.0	908	11 CNS06105	AL400539 T3 end of
9	100	69.0	2045	8 SCSDMNG	X02156 Yeast gene
10	100	69.0	18715	8 YSCHL2825	U10400 Saccharomyc
11	100	69.0	24431	8 YSCH9780	U0555 Saccharomyc
12	99	68.3	1071	1 AF061031	AF061031 Mycobacte
13	99	68.3	1478	1 MFSODA	X70914 M. fortitulum
14	98	67.6	649	1 MLEPSOD	X16453 M. fortitulum
15	98	67.6	344050	1 MLEPSOD	AL583917 Mycobacte
16	94	64.8	790	1 MTSD	X52861 M. tuberculosis
17	94	64.8	943	1 AF077406	AF077406 Mycobacte
18	94	64.8	1321	1 AF061030	AF061030 Mycobacte
19	94	64.8	14282	1 AB007188	AB007188 Mycobacte
20	91	64.8	37751	1 MTCY1A6	X91650 P. freudentre
21	91	62.8	567	1 PFSOD	Y09012 P. freudentre
22	91	62.8	567	1 PFSOD	AB079877 Marsupena
23	80	55.2	655	3 AB079877	AB079877 Marsupena
24	80	55.2	954	4 AB001693	AB001693 Equus cab
25	80	55.2	4439	1 GSP312188	AJ312188 Gordonia
26	79	54.5	812	6 E15569	E15569 cDNA encodi
27	77	53.1	600	6 AX123311	AX123311 Sequence
28	77	53.1	606	4 RABMSD	I28808 Oryctolagus
29	77	53.1	960	4 AB055218	AB055218 Corynebac
30	77	53.1	1143	6 AX136077	AX136077 Sequence
31	77	53.1	1143	6 BD011434	BD011434 DNA, amin
32	77	53.1	2207	1 AF236111	AF236111 Corynebac
33	77	53.1	309400	6 AX127153	AX127153 Sequence
34	77	53.1	325651	1 AP005283	AP005283 Corynebac
35	76	52.4	663	3 AF521909	AF521909 Trichinel
36	76	52.4	761	8 DLA278864	AJ278864 Digitalis
37	74	51.0	325	6 AX469763	AX469763 Sequence
38	74	51.0	535	10 S7883281	S7883281 Sod-2-manga
39	74	51.0	594	6 ARI30379	ARI30379 Sequence
40	74	51.0	600	6 A12180	A12180 Artificial
41	74	51.0	600	6 A12190	A12190 Artificial
42	74	51.0	600	6 A12191	A12191 Artificial m
43	74	51.0	600	6 A12200	A12200 Nucleotide
44	74	51.0	609	6 E03557	E03557 Synthetic D
45	74	51.0	636	1 AF236110	AF236110 Acidianus

## ALIGNMENTS

RESULT 1

AF031478 1859 bp DNA linear PLN 13-JAN-2000  
LOCUS AF031478  
DEFINITION Candida albicans manganese-superoxide dismutase precursor (SOD2)  
gene, complete cds.  
AF031478  
VERSION AF031478.1 GI:2623884  
KEYWORDS  
SOURCE  
ORGANISM  
Candida albicans.  
Candida albicans  
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
Saccharomycetales; mitosporic Saccharomycetales; Candida.  
REFERENCE  
1 (bases 1 to 1859)  
Rhee,G.E., Hwang,C.S., Brady,M.J., Kim,S.T., Kim,Y.R., Huh,W.K.,  
Baek,Y.U., Lee,B.H., Lee,J.S. and Kang,S.O.  
Manganese-containing superoxide dismutase and its gene from Candida  
albicans  
JOURNAL Biochim. Biophys. Acta 1426 (3), 409-419 (1999)  
MEDLINE 99177423  
JOURNAL 10076057  
PUBMED 2 (bases 1 to 1859)  
REFERENCE Kang,S.O. and Rhee,G.  
Direct Submission  
Submitted (26-OCT-1997) Microbiology, Seoul National University,  
Shinlim-dong, Kwanak-gu, Seoul 151-742, Republic of Korea  
JOURNAL  
TITLE  
AUTHORS  
FEATURES  
Location/Qualifiers  
1. 1859  
/organism="Candida albicans"  
/db\_xref="taxon:5476"  
gene 848. 1552  
/gene="SOD2"  
848. 1552  
/gene="SOD2"  
/note="Mn-SOD"  
/codon\_start=1  
/transl\_table=12  
/product="manganese-superoxide dismutase precursor"  
/protein\_id="AAB86583.1"  
/db\_xref="GI:2623885"  
/translation="MFSSIRSSRVLLKSSATTRATLNAASKTFRKYSPLDYR  
FSATRPYSGINRIHYTKHQTYYNNINAEIQAQVKGKGVKLVAKLNPNP  
GGYINHCILMKNLAPVSGQPSQSDSKLGRQIVKQFSLDKLLETNGKLAGIQSS  
GWAIVKKNKANGDITDVTITANQDITVDPLPLAIDAWEHAYYLOYNKADYFN  
LMHYINMKAEARREF"  
848. 949  
/gene="SOD2"  
950. 1549  
/gene="SOD2"  
/product="manganese-superoxide dismutase"  
BASE COUNT 601 a 310 c 312 g 636 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.44e-15 Length: 1859  
Score: 141.00 Matches: 27  
Percent Similarity: 96.43% Conservative: 0  
Best Local Similarity: 96.43% Mismatches: 1  
Query Match: 97.24% Indels: 0  
Gaps: 0  
US-09-987-190-2 (1-30) x AF031478 (1-1859)  
Qy 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20  
Db 950 AATATATGTTTCCAGAAATGACATATTCCTCGCTACTGAACCATACATTTCTGCT 1009  
Qy 21 GlnIleAsnGluIle\*\*TyrThr 28  
Db 1010 CAATAAACGAATTCACATCACT 1033  
RESULT 2  
LOCUS CSNNNSOD 2332 bp DNA linear PLN 11-MAY-2001  
DEFINITION Candida sp. HN95 MnSOD gene.

Y11598  
VERSION Y11598.1 GI:1877052  
KEYWORDS manganese superoxide dismutase; MnSOD gene.  
SOURCE  
ORGANISM  
Candida sp. HN95.  
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
Saccharomycetales; mitosporic Saccharomycetales; Candida.  
REFERENCE  
1 (bases 1 to 2332)  
Hong,Y.M., Nam,Y.S. and Choi,S.Y.  
AUTHORS  
TITLE  
JOURNAL Molecular Cloning and Characterization of Mn-Superoxide Dismutase  
Gene from Candida sp  
J. Microbiol. 35, 309-314 (1997)  
REFERENCE 2 (bases 1 to 2332)  
AUTHORS Choi,S.  
TITLE Direct Submission  
JOURNAL Submitted (04-MAR-1997) S. Choi, Hannam University, Microbiology,  
133 Ojung-dong, Taejeuk-ku, Taejeon, 300-791, SOUTH KOREA  
FEATURES  
Location/Qualifiers  
1. 2332  
/organism="Candida sp. HN95"  
/strain="HN95"  
/db\_xref="taxon:159257"  
1435. 2332  
/gene="MnSOD"  
1435. 2332  
/gene="MnSOD"  
1500. 2183  
/gene="MnSOD"  
/gene="MnSOD"  
/codon\_start=1  
/protein\_id="CAA72335.1"  
/db\_xref="SPTREMBL:P79022"  
/db\_xref="GI:2398587"  
/translation="MLSSAIKRSVAVGARSVSSVGAVRTKSLPDLPMDFGALP  
HISGQINRIHYTKHQTYYNNINAEIQAQVKGKGVKLVAKLNPNP  
LPMKNLAPKQGGGPRPAEDSERAKRIVEYGSGLDNKATITNGKLAGIQSSGNAFYIK  
NKEGGLADIVITANQDITVDLPVPLAIDAWEHAYYLOYNKADYFKAIWVYINMK  
EAERKRYLIN"  
1500. 1589  
/gene="MnSOD"  
1590. 2180  
/gene="MnSOD"  
/product="unnamed"  
BASE COUNT 646 a 519 c 442 g 725 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 2.86e-09 Length: 2332  
Score: 107.00 Matches: 20  
Percent Similarity: 85.71% Conservative: 4  
Best Local Similarity: 71.43% Mismatches: 4  
Query Match: 73.79% Indels: 0  
Gaps: 0  
US-09-987-190-2 (1-30) x CSNNNSOD (1-2332)  
Qy 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20  
Db 1584 AAAGTATGTTTCCAGAAATGACATTCCTCGAGCTTGAGACCTCACAATTCGGGA 1643  
Qy 21 GlnIleAsnGluIle\*\*TyrThr 28  
Db 1644 CAGATCAACGAATTCACATCAACC 1667  
RESULT 3  
LOCUS NAO02341 735 bp DNA linear BCT 12-MAR-1996  
DEFINITION Nocardia asteroides GUH2 superoxide dismutase gene, complete cds.  
ACCESSION U02341  
VERSION U02341.1 GI:484066  
KEYWORDS  
SOURCE  
ORGANISM  
Nocardia asteroides.  
Nocardia asteroides  
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;



REFERENCE 1 Actinomycetales; Corynebacterineae; Nocardiaceae; Nocardia.  
 AUTHORS 1 (bases 1 to 735)  
 TITLE Alcedor, D.J., Chapman, G.D. and Beaman, B.L.  
 JOURNAL Isolation, sequencing and expression of the superoxide  
 MEDLINE dismutase-encoding gene (sod) of Nocardia asteroides strain GUN-2  
 PUBMED Gene 164 (1), 143-147 (1995)  
 7590304  
 2 (bases 1 to 735)  
 REFERENCE Chapman, G.G.  
 AUTHORS Direct Submission  
 JOURNAL Submitted (01-OCT-1993) Gail D. Chapman, Medical Microbiology and  
 Immunology, University of California at Davis, Medical Science  
 Building 1A, Davis, CA 95616, USA  
 location/Qualifiers  
 FEATURES  
 source 1..735  
 /organism="Nocardia asteroides"  
 /strain="GUN2"  
 /db\_xref="taxon:1824"  
 /clone="clone PAC1500"  
 /clone\_11b="Library EMBL-3"  
 99..106  
 /note="potential ribosome binding site"  
 112..735  
 /EC\_number="1.15.1.1"  
 /standard\_name="SOD"  
 /note="alternative start codon GTG"  
 /codon\_start=1  
 /transl\_table=11  
 /product="superoxide dismutase"  
 /protein\_id="AA01964.1"  
 /db\_xref="GI:484067"  
 /translation="MAEYTLPLDWDYALPEPHISQGINELHSHKHAAYVAGANTL  
 EKLEBARANEDHAIPLNEKNLAFHLEGHVNSIMWNLSPPGDKPVGELAAIDDA  
 FGSFDRRAQFSAANGLOGSGMAVGYDPLVGSRLTLTLFOLYDQANVPLGIIPLLQVD  
 MMEHAFILQIKNKYADYVKAFAFNVNADYQKRYAATSAQGLIFG"  
 BASE COUNT 148 a 242 c 230 g 115 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 3.64e-09 Length: 735  
 Score: 103.00 Matches: 17  
 Percent Similarity: 92.86% Conservative: 9  
 Best Local Similarity: 60.71% Mismatches: 2  
 Query Match: 71.03% Indels: 0  
 Gaps: 0  
 DB: 1  
 9-987-190-2 (1-30) x NAU02341 (1-735)  
 QY 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20  
 Db 118 GAGTACACGCGCGGATTCGATTACGACTACACGCGCCCGAACCCACATCTCCGGG 177  
 QY 21 GlnIleAsnGluIle\*\*\*TyrThr 28  
 Db 178 CAGATCAACGAGCTGCACATTCC 201  
 RESULT 4  
 AF333434 AF333434 666 bp DNA linear BCT 17-JAN-2001  
 LOCUS Mycobacterium avium subsp. paratuberculosis superoxide dismutase  
 DEFINITION gene, complete cds.  
 ACCESSION AF333434  
 VERSION AF333434  
 KEYWORDS AF333434.1 GI:12247898  
 SOURCE Mycobacterium avium subsp. paratuberculosis.  
 ORGANISM Mycobacterium avium subsp. paratuberculosis  
 Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 Actinomycetales; Corynebacterineae; Mycobacteriaceae;  
 Mycobacterium; Mycobacterium avium complex (MAC).  
 1 (bases 1 to 666)  
 REFERENCE Dheenadhayalan, V. and Chang, Y.F.  
 AUTHORS Mycobacterium avium subsp. paratuberculosis superoxide dismutase  
 TITLE

JOURNAL gene  
 REFERENCE 2 Unpublished  
 AUTHORS 2 (bases 1 to 666)  
 TITLE Dheenadhayalan, V. and Chang, Y.F.  
 JOURNAL Direct Submission  
 Submitted (04-JAN-2001) Population Medicine and Diagnostic Science,  
 College of Veterinary Medicine, Cornell University, P.O. Box 5786,  
 Ithaca, NY 14853, USA  
 location/Qualifiers  
 FEATURES  
 source 1..666  
 /organism="Mycobacterium avium subsp. paratuberculosis"  
 /sub\_species="paratuberculosis"  
 /db\_xref="taxon:1770"  
 34..666  
 /codon\_start=1  
 /transl\_table=11  
 /product="superoxide dismutase"  
 /protein\_id="AA03084.1"  
 /db\_xref="GI:12247899"  
 /translation="MAEYTLPLDWDYALPEPHISQGINELHSHKHAAYVAGANTL  
 AKLEBARANEDHAIPLNEKNLAFHLEGHVNSIMWNLSPPGDKPVGELAAIDDA  
 FGSFDRRAQFSAANGLOGSGMAVGYDPLVGSRLTLTLFOLYDQANVPLGIIPLL  
 QVDMMEHAFILQIKNKYADYVKAFAFNVNADYQKRYAATSAQGLIFG"  
 BASE COUNT 142 a 221 c 196 g 107 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 7.44e-09 Length: 666  
 Score: 101.00 Matches: 17  
 Percent Similarity: 92.86% Conservative: 9  
 Best Local Similarity: 60.71% Mismatches: 2  
 Query Match: 69.66% Indels: 0  
 Gaps: 0  
 DB: 1  
 US-09-987-190-2 (1-30) x AF333434 (1-666)  
 QY 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20  
 Db 40 GATACACCGCTGCGCGACTGAGCTACGAGCTGTTGGAACCGCACATCTCCGGG 99  
 QY 21 GlnIleAsnGluIle\*\*\*TyrThr 28  
 Db 100 CAGATCAACGAGTCCACACACC 123  
 RESULT 5  
 MAU11550 MAU11550 960 bp DNA linear BCT 25-SEP-1996  
 LOCUS Mycobacterium avium TMC 724 superoxide dismutase (sod) gene,  
 complete cds.  
 ACCESSION U11550  
 VERSION U11550.1 GI:555745  
 KEYWORDS  
 SOURCE Mycobacterium avium.  
 ORGANISM Mycobacterium avium  
 Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 Actinomycetales; Corynebacterineae; Mycobacteriaceae;  
 Mycobacterium; Mycobacterium avium complex (MAC).  
 1 (bases 1 to 960)  
 REFERENCE Bescuyer, V., Haddad, N., Frehel, C. and Berche, P.  
 AUTHORS Molecular characterization of a surface-exposed superoxide  
 TITLE dismutase of Mycobacterium avium  
 JOURNAL Microb. Pathog. 20 (1), 41-55 (1996)  
 MEDLINE 96276149  
 PUBMED 8692009  
 2 (bases 1 to 960)  
 REFERENCE Bescuyer, V.E.  
 AUTHORS Direct Submission  
 JOURNAL Submitted (28-JUN-1994) Vincent E. Bescuyer, INSERM  
 U411-Microbiologie, Faculte de Medecine Necker, 156 rue de  
 Vaugirard, Paris 75015, France  
 location/Qualifiers  
 FEATURES  
 source 1..960  
 /organism="Mycobacterium avium"

gene  
/strain="TWC 724"  
/db\_xref="taxon:1764"  
/clone="psod1"  
CDS  
251..874  
/gene="bod"  
251..874  
/gene="bod"  
/function="molecular dimutation of oxygen free radicals"  
/note="GTG start codon"  
/codon\_start=1  
/evidence=experimental  
/transl\_table=1  
/product="superoxide dismutase"  
/protein\_id="AA08770.1"  
/db\_xref="GI:555746"  
/translation="MAEYTLPDLDMDVYALPEPHISQINIEHTKHATYKGVNDALAKLEBARANEDHAIFLNEKNLAFHGHVNHSTWKNLSPDGDKFTGEILAAIDDAFGSFDKRAQFSANANGLOSGMVLGIDTIGSRLLTFQYLDQANVPLGITPLLOYDMMHAFYLYQKNVADYVKAFMNVVNMADYQKRYAATSKAOGILFG"

BASE COUNT 179 a 320 c 298 g 163 t

ORIGIN

Alignment Scores:  
Pred. No.: 1,17e-08 Length: 960  
Score: 101.00 Matches: 17  
Percent Similarity: 92.86% Conservative: 9  
Best Local Similarity: 60.71% Mismatches: 2  
Query Match: 69.66% Indels: 0  
DB: 1 Gaps: 0

US-09-987-190-2 (1-30) x MAU11550 (1-960)

Qy 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20  
Db 257 GAATACACCTGCGCCGACTGGACTATGACGCTTGGAACCGACATCTCGGG 316  
|||  
Qy 21 GlnIleAngIuIle\*\*TyrThr 28  
|||  
Db 317 CAGATCAACGAGATCCACACACC 340  
|||

RESULT 6  
LOCUS AF180816 1741 bp DNA linear BCT 30-NOV-2001  
DEFINITION Mycobacterium avium subsp. paratuberculosis superoxide dismutase (sod) gene, complete cds; and unknown gene.  
ACCESSION AF180816.1 GI:9957589  
VERSION AF180816.1  
KEYWORDS  
SOURCE Mycobacterium avium subsp. paratuberculosis.  
ORGANISM Mycobacterium avium subsp. paratuberculosis  
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium avium complex (MAC).  
KEYWORDS 1 (bases 1 to 1741)  
SOURCE Liu, X., Feng, Z., Harris, N.B., Cirillo, J.D., Bercovier, H. and Barletta, R.G.  
REFERENCE Identification of a secreted superoxide dismutase in Mycobacterium avium ssp. paratuberculosis  
JOURNAL FEMS Microbiol. Lett. 202 (2), 233-238 (2001)  
MEDLINE 21411748  
PUBMED 11520620  
2 (bases 1 to 1741)  
AUTHORS Liu, X., Feng, Z., Cirillo, J. and Barletta, R.G.  
TITLE Direct Submission  
JOURNAL Submitted (25-AUG-1999) Veterinary and Biomedical Sciences, University of Nebraska-Lincoln, Fair Street and East Campus Loop, Lincoln, NE 68583-0905, USA  
FEATURES  
source Location/Qualifiers  
1..1741  
/organism="Mycobacterium avium subsp. paratuberculosis"  
/strain="K-10"  
/db\_species="paratuberculosis"  
/db\_xref="taxon:1770"

gene  
260..883  
/gene="sod"  
260..883  
CDS  
/gene="sod"  
/codon\_start=1  
/transl\_table=1  
/product="superoxide dismutase"  
/protein\_id="AAG09425.1"  
/db\_xref="GI:9957590"  
/translation="MAEYTLPDLDMDVYALPEPHISQINIEHTKHATYKGVNDALAKLEBARANEDHAIFLNEKNLAFHGHVNHSTWKNLSPDGDKFTGEILAAIDDAFGSFDKRAQFSANANGLOSGMVLGIDTIGSRLLTFQYLDQANVPLGITPLLOYDMMHAFYLYQKNVADYVKAFMNVVNMADYQKRYAATSKAOGILFG"  
1041..1574  
/note="similar to Mycobacterium tuberculosis hypothetical protein Rv3847"  
/codon\_start=1  
/transl\_table=1  
/product="unknown"  
/protein\_id="AAG09426.1"  
/db\_xref="GI:9957591"  
/translation="MENGSDRPIGVGPFHSGALKGPIVSGRPDSTKEMALIMVAVRVASLRLSTTVFGARRELPDPEPGITGVLAGSTVGEAIPGHADHPRLLMHAPSETPSPSECTGAASGCVLPGLPYLLEHRAVVEADGITSMVSRGVDPISHDPTALIMLAA"

BASE COUNT 286 a 594 c 562 g 299 t

ORIGIN

Alignment Scores:  
Pred. No.: 2,45e-08 Length: 1741  
Score: 101.00 Matches: 17  
Percent Similarity: 92.86% Conservative: 9  
Best Local Similarity: 60.71% Mismatches: 2  
Query Match: 69.66% Indels: 0  
DB: 1 Gaps: 0

US-09-987-190-2 (1-30) x AF180816 (1-1741)

Qy 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20  
Db 266 GAATACACCTGCGCCGACTGGACTATGACGCTTGGAACCGACATCTCGGG 325  
|||  
Qy 21 GlnIleAngIuIle\*\*TyrThr 28  
|||  
Db 326 CAGATCAACGAGATCCACACACC 349  
|||

RESULT 7  
LOCUS D13288 815 bp DNA linear BCT 17-JUN-1998  
DEFINITION Mycobacterium lepraemurium DNA for Mn superoxide dismutase, complete cds.  
ACCESSION D13288  
VERSION D13288.1 GI:3228263  
KEYWORDS Mn superoxide dismutase.  
SOURCE Mycobacterium lepraemurium (strain:Hawaiian) DNA, clone\_l1b:lambda gt 11 and pUC118 clone:pMB67 and pMP49.  
ORGANISM Mycobacterium lepraemurium  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
KEYWORDS 1 (bases 1 to 815)  
SOURCE Nakamura, M.  
REFERENCE Unpublished  
2 (bases 1 to 815)  
AUTHORS Nakamura, M.  
TITLE Direct Submission  
JOURNAL Submitted (21-SEP-1992) Masahiko Nakamura, Osaka University, 565, Institute for Protein Research, 3-2 Yamadaoka, Suita, Osaka, Japan (E-mail:masahiko@protein.osaka-u.ac.jp, Tel:06-879-8628, Fax:06-879-8629)  
FEATURES  
source Location/Qualifiers  
1..815  
/organism="Mycobacterium lepraemurium"  
/strain="Hawaiian"



```

/organism="Saccharomyces cerevisiae"
/strain="FL 100"
/db_xref="taxon:4932"
74..271
/note="URF 1"
/codon_start=1
/protein_id="CAA26091.1"
/db_xref="GI:600040"
/db_xref="SPTREMBL:007073"
/translation="MKKFMHRAVLFYLFKLPBPNNYIFNEYLNLYISFFFLSLIS
LFLFLSWRLSLPLRMVRLHY"
218..239
/note="put. upstream activation site (UAS) 1 involved in
reg. of MnsOD expression"
311..329
/note="put. UAS2"
559..1260
/gene="mnsod"
559..1260
/gene="mnsod"
/codon_start=1
/protein_id="CAA26092.1"
/db_xref="GI:4514"
/db_xref="SWISS-PROT:P00447"
/translation="MFAKTAANLTKKGLSLSTTARTKVTLPDLKWDGALPEYI
SGQINELHYTHQTYNGFNTAYDQFELSDDLAKPSPANARMTAIDONIKFHGG
GFTNHCLEFWMENLAPESQGGSEPTGALKAIDEQGSIDELIKLTNTYLAQVQSSGMA
FIVKRLSNGKLDVQYNTQDVTGPIVLVAIDAMEHAYVLYQNKRADYFKALMNV
VNMEASRRPDAGKI"
559..639
/gene="mnsod"
640..1257
/gene="mnsod"
/product="unnamed"
1384..1599
/note="URF 2"
/codon_start=1
/protein_id="CAA26093.1"
/db_xref="GI:4517"
/db_xref="SPTREMBL:007074"
/translation="MHRKRKKKCKRTEKDNNTNLPFFFPFCSLSPTLLAPHYIIP
TRLTHQAEHQFLFLFQPIIVKPLRS"
BASE COUNT 622 a 497 c 384 g 541 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 4,54e-08 Length: 2045
Score: 100.00 Matches: 18
Percent Similarity: 82.14% Conservative: 5
Local Similarity: 64.29% Mismatches: 5
Indels: 0
Gap: 68.97% Gaps: 0
DB:
US-09-987-190-2 (1-30) x SCSODMNG (1-2045)
QY 1 lveTySerleuprogileuagpyrGluPheserAlathrgupProtyrIleSergly 20
Db 637 AAGGTACCTTCGACAGCTTGAAGGACTTCGGTCACTGGAACCTTATATCTCCGT 656
QY 21 GlnIleAsnGluIle**TyrtYr 28
Db 697 CAATCAGCAATTGCATTACACC 720
RESULT 10
LOCUS YSCHL2825 18715 bp DNA linear PLN 04-SEP-1997
DEFINITION Saccharomyces cerevisiae chromosome VIII cosmid L2825.
ACCESSION U10400 U00093
VERSION U10400.1 GI:500701
KEYWORDS
SOURCE Saccharomyces cerevisiae.
ORGANISM Saccharomyces cerevisiae
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

```

```

REFERENCE
AUTHORS
1 (bases 1 to 18715)
Johnston, M., Andrews, S., Brinkman, R., Cooper, J., Ding, H., Dover, J.,
Du, Z., Favello, A., Fulton, L., Gattung, S., Geisel, C., Kirsten, J.,
Kucaba, T., Hillier, L., Jier, M., Johnston, L., Kepler, D.,
Langston, Y., Latreille, P., Louis, E., Macri, C., Mardis, E.,
Mouser, L., Nhan, M., Rifken, L., Riles, L., St. Peter, H., Thornton, L.,
Trevaskis, E., Vaudin, M., Vaughan, K., Vignati, D., Wilcox, L.,
Wills, A., Wilson, R., Wohlman, P. and Waterston, R.
Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
VIII
Science 265 (5181), 2077-2082 (1994)
JOURNAL
MEDLINE
94378003
PUBMED
8091229
REFERENCE
2 (bases 1 to 18715)
Du, Z.
The sequence of S. cerevisiae cosmid L2825
Unpublished (1994)
3 (bases 1 to 18715)
Waterston, R.
Direct Submission
Submitted (06-JUN-1994)
4 (bases 1 to 18715)
Jia, Y. and Cherry, J.M.
Direct Submission
Submitted (04-SEP-1997)
JOURNAL
REFERENCE
AUTHORS
JOURNAL
TITLE
JOURNAL

```

## COMMENT

Curated by:  
 Saccharomyces Genome Database  
 URL: <http://genome-www.stanford.edu/>  
 e-mail: [yeast-curator@genome.stanford.edu](mailto:yeast-curator@genome.stanford.edu)

Submitted by:  
 Genome Sequencing Center  
 Department of Genetics, Washington University,  
 St. Louis, MO 63110, USA  
 e-mail: [mj@sequencer.wustl.edu](mailto:mj@sequencer.wustl.edu)

## NEIGHBORING COSMID INFORMATION:

This sequence includes nucleotides 1-17118 of lambda clone  
 YSCH2825. It overlaps with the cosmid on the left (YSCH9780) by  
 800 bp; with the cosmid on the right (YSCH8082) by 1001 bp.

## FEATURES

## source

```

1..18715
/organism="Saccharomyces cerevisiae"
/strain="S288C (AB972)"
/db_xref="taxon:4932"
/map="VIII"
complement(128..829)
/gene="SOD2"
complement(128..829)
/gene="SOD2"
complement(128..829)
/gene="SOD2"
complement(128..829)
/note="YHR080C"
/codon_start=1
/product="Sod2p: Superoxidase dismutase"
/protein_id="AAB68939.1"
/db_xref="GI:500704"
/db_xref="SPTREMBL:007074"
/translation="MFAKTAANLTKKGLSLSTTARTKVTLPDLKWDGALPEYI
SGQINELHYTHQTYNGFNTAYDQFELSDDLAKPSPANARMTAIDONIKFHGG
GFTNHCLEFWMENLAPESQGGSEPTGALKAIDEQGSIDELIKLTNTYLAQVQSSGMA
FIVKRLSNGKLDVQYNTQDVTGPIVLVAIDAMEHAYVLYQNKRADYFKALMNV
VNMEASRRPDAGKI"
complement(1348..2904)
/gene="YHR099C"
complement(1348..2904)
/gene="YHR099C"
complement(1348..2904)
/codon_start=1
/evidence="not experimental"
/product="Yhr099cp"

```

gene  
CDS  
/protein\_id="AAB68938.1"  
/db\_xref="GI:500703"  
/translation="MHPPTQTPSKNSGKHHIVVGGIIGCCTAYVLTQHPSPSP  
THHTITLESRIAGSAGKAGLLASVNPPIQVLSFQHELSDEGNSNMNDYR  
LTVSLADVLEEYIENYERLAKVYANVPPKRPPIYISNKNFNGISNLSGSS  
SLKNDASNEEGSDIVSSVPSLHSLTNERMSHMSASDLSVSVEQLRETNH  
NPLPADDMIRRELVNDMSLGGITDTTRQQLPIYFTHFLSKAMETGAVDLLGVG  
LKCEMDCVHSLKLPVSVKNRNSRGAENPDIKLTGFENDENAKPIEINDIQIVL  
SMGPWTSKILNDCEPISGLRAHSVTKSEKESVSPALIAELKNDREFSPENYAKD  
EVVYCGSDTLVNIPESSDVEVSEKDELYHVSRLPSLKGHLRKAQCFPLV  
NPTSSGPIIGETVNDKLYINSGHSCWGINNAPATGKLMETLIDGELTSLSLDP  
KLYPDATILS"  
3759..4730  
/gene="Rpl27"  
join(3759..4730,4351..4730)  
/gene="Rpl27"  
/note="YHR010W"  
/codon\_start=1  
/evidence=not\_experimental  
/product="Rpl27p: Probable 60S ribosomal protein l27"  
/protein\_id="AAB68944.1"  
/db\_xref="GI:500709"  
/translation="MAKPLKAGKAVVVRGVRGAKVVIYVKNHSGSKSPFGHALV  
GIRYPIKATYKKGAKKAVAKRTKIKPKIKVNNHLPTRITLIDVFAKSVSTETTE  
QPSGEAKKVVAKFAFERHQAGKNQWFFSLRF"  
5018..6358  
/gene="YHR011W"  
5018..6358  
/gene="YHR011W"  
/note="Similar to Seryl-tRNA synthetase"  
/codon\_start=1  
/evidence=not\_experimental  
/product="Yhr011wp"  
/protein\_id="AAB68940.1"  
/db\_xref="GI:500705"  
/translation="MITIRLFSISNRSEFLKPPQDVYKTIEMIPQYOTSQNFELIE  
ADSLIRSLQELGERYQNIKEIDKVIADIQIQRSIEAQIKDKRTITEYSALALKE  
QYNDQNSKSELEKRILETCKSLPNTLDPLVLPALPOIEQWNPINPKTHKTSQVAPND  
IMLKNNMLDQATNSVNGSWYLLNDGARLEQALVAVALKKANENGSSCVSPSITK  
KELIDACGFNPDNNRROIYALQDTNIGLVATAEIPLAGANKVLEINSGESKLV  
GVSRVCYRAEGARGKDKTGLYVHEFTKYVLFPCMSKPEYSAKYLEELKQIQISVVE  
LGIPAKYLMPSNDLGNPAFKKDIIEAMPGRGKFGELISSASNLCTDFOSRLNTKYRD  
DNTEKLYHTLTSTAMAIPRIVAVLENFNPSTGKISVECELRFPNNGORYI"  
6719..7686  
/gene="VPS29"  
join(6719..7686,6886..7686)  
/gene="VPS29"  
/note="YHR012W"  
/codon\_start=1  
/evidence=not\_experimental  
/product="Vps29p: Protein involved in vacuolar protein  
sorting"  
/protein\_id="AAB68947.1"  
/db\_xref="GI:500712"  
/translation="MLILALSDAHIIPRATDLPVKFKLLSVDPKISQVALLGNSTKS  
YDFLKFVNOISNNITIVRGEDNGHLPSYTKDKASDNRSEIIPMSIIRQGLKIG  
CCSGYIVPKNPDSLILAROLDVDLITLNGTNAVAYTLGEGFVNPSCGTAFPT  
DMPIVPEVDSDEAVTSEVOKPTKENOSDDADGSGTGKQPSGYSYTPKSTAGERRN  
ENESNVKPEQKEDVDMDSIDNGSNSBSFCLLDIQNTCTIYIILYNGEVKVDK  
VYVEKE"  
complement(7968..8684)  
/gene="ARD1"  
complement(7968..8684)  
/gene="ARD1"  
/note="YHR013C"  
/codon\_start=1  
/product="Ard1p: subunit of the major N  
alpha-acetyltransferase"  
/protein\_id="AAB68937.1"  
/db\_xref="GI:500702"  
/translation="MPINIRATINDIICQNNANLHNPENYMKYTMHLSPPAS  
FVATTTLDCEDSDEQDNDKLTLLTGNDGRTIKLDPYTLAAGEKLVGVTLVKNAD  
PDQQNEPENGHTLSLVMTYRRMGIAENIMRQALPALREVHAQAEVYSLHVRQSNRA

ALHLYRDTLAEVLISIEKSYQDGEDAYAMKVKIKLEELQISNFTHRLKENEKLED  
DLSEDLLEDIIKQGVNDIIV"  
9284..10159  
/gene="SPO13"  
9284..10159  
/gene="SPO13"  
/note="YHR014W"  
/codon\_start=1  
/product="Spo13p: Meiosis-specific protein"  
/protein\_id="AAB68941.1"  
/db\_xref="GI:500706"  
/translation="MAPRKPRLLELSPHSHKRVQKPLQEKTPNLRVSPFLAKIGK  
EIKKKEIRKTKTESNI FNSKHVDLLESHPGLNVSQAQYSSKAGDVRVLYNKS  
NTLNGERQTERPFSNLSRPFEDIEQPKSTFVLSSQINVEREAPMPVPYXIA  
PSPMYNFSPONFVGNPTPLPSHNPINAYAIPIOREPLYPNVNVYDSPFKKTRLP  
HOTSLDKENYQVLPYVPSISNNGDFVGEPTFRAPKLSKRLSTLVDNCSYDS  
SGQNAVYNDSESSIN"  
complement(10264..10345)  
/gene="ts(AGA)H"  
complement(10264..10345)  
/gene="ts(AGA)H"  
/product="tRNA-Ser"  
/note="codon recognized: UCU"  
10911..11246  
/note="delta element"  
11559..11630  
/gene="tQ(UUG)H"  
11559..11630  
/gene="tQ(UUG)H"  
/product="tRNA-Gln"  
/note="codon recognized: CAA"  
11791..13770  
/gene="YHR015W"  
11791..13770  
/gene="YHR015W"  
/note="Similar to polyadenylate-binding protein in  
N-terminus."  
/codon\_start=1  
/evidence=not\_experimental  
/product="Yhr015wp"  
/protein\_id="AAB68942.1"  
/db\_xref="GI:500707"  
/translation="MPNSHGNVNNISLNSKONRISISKSCPNDKAROKSFKTSIQ  
ALVAVQAGYKLDGDKDAVEKENSILKKYDCKNATQKKEQOEVEKTVAKGSVOK  
YITKTSNLSLFGNLKSTYELKRIPIRYSFSAKVCRDPLTKKISGYGLNPK  
DKNDASARKFNTYVFGQVAKIMPSKNLIRKNIIGTVFNSNPLENPQLTTRSF

Alignment Scores:  
Pred. No.: 7.07e-07 Length: 18715  
Score: 100.00 Matches: 18  
Percent Similarity: 82.14% Conservative: 5  
Best Local Similarity: 64.29% Mismatches: 5  
Query Match: 68.97% Indels: 0  
Gaps: 0

US-09-987-190-2 (1-30) x YSCHL2825 (1-18715)  
OY 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20  
Db 751 AAAGTCACCTTGCAGACTTGAAGTGGACCTTCGTGCATTGAACCTTATATCTCCGGT 692  
OY 21 GlnIleAngIuile\*\*\*TyrThr 28  
Db 691 CAAATCAACGAATTGCATTACACC 668

RESULT 11  
YSCH9780/c 24431 bp DNA linear PLN 03-SEP-1997  
LOCUS YSCH9780  
DEFINITION Saccharomyces cerevisiae chromosome VIII cosmid 9780.  
ACCESSION U10555 U00093  
VERSION U10555.1 GI:500813  
KEYWORDS  
SOURCE Saccharomyces cerevisiae.

ORGANISM Saccharomyces cerevisiae  
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
Saccharomycetalei; Saccharomycetaceae; Saccharomycetes.  
REFERENCE 1 (bases 1 to 24431)  
AUTHORS Johnston, M., Andrews, S., Brinkman, R., Cooper, J., Ding, H., Dover, J.,  
Du, Z., Favelli, A., Fulton, L., Gattung, S., Gelsel, C., Kisten, J.,  
Kusaba, T., Hillier, L., Jier, M., Johnston, L., Keppler, D.,  
Kugan, Y., Latreille, P., Louis, E., Macri, C., Mardis, E.,  
Moussier, L., Nhan, M., Rife, L., Riles, L., St. Peter, H., Thornton, L.,  
Tverskoy, E., Vaudin, M., Vaughan, K., Vignati, D., Wilcox, L.,  
Wills, A., Wilson, R., Woldman, P. and Waterston, R.  
TITLE Complete nucleotide sequence of Saccharomyces cerevisiae chromosome  
VIII  
JOURNAL Science 265 (5181), 2077-2082 (1994)  
MEDLINE 94378003  
PUBMED 8091229  
REFERENCE 2 (bases 1 to 24431)  
AUTHORS Favelli, T.  
TITLE The sequence of S. cerevisiae cosmid 9780  
JOURNAL Unpublished (1994)  
AUTHORS 3 (bases 1 to 24431)  
TITLE Waterston, R.  
JOURNAL Direct Submission  
AUTHORS Submitted (10-JUN-1994)  
JOURNAL 4 (bases 1 to 24431)  
AUTHORS Jia, Y. and Cherry, J. M.  
TITLE Direct Submission  
JOURNAL Submitted (03-SEP-1997)  
COMMENT Curated by:  
Saccharomyces Genome Database  
URL: http://genome-www.stanford.edu/  
e-mail: yeast-curator@genome.stanford.edu  
Submitted by:  
Genome Sequencing Center  
Department of Genetics, Washington University,  
St. Louis, MO 63110, USA  
e-mail: mj@sequencer.wustl.edu  
NEIGHBORING COSMID INFORMATION:  
This sequence includes nucleotides 10590-34824 of cosmid 9780. The  
adjacent clone on the right is YSCHL2825. The adjacent clone on  
the left is YSCHL5018.  
location/Qualifiers  
1. 24431  
/organism="Saccharomyces cerevisiae"  
/strain="S288C (AB972)"  
/db\_xref="taxon:4932"  
/map="VIII"  
90..1274  
/gene="MRP4"  
90..1274  
/gene="MRP4"  
/note="YHL004W"  
/codon\_start=1  
/product="Mrp4: Mitochondrial ribosomal protein"  
/protein\_id="AAB68428.1"  
/db\_xref="GI:500819"  
/translation="MQRVFARNFRRLNSLTPSTKRFQSSASAGANTPNNDEVM  
LQQLKLYDEIRSELKSLSQVPEDELPELKLSDKQLEQLLEALSPFRYAL  
LNKLFDSTKLDGQSTTTAAATPTKYPNLLPSANDKEYSQELFLRLNHSMTAKL  
GATISKVYYPKHDIFYPPLPENITVESLMSAGVHGOSTLSRSTOSYIAGEYGIH  
IIDNQTLSTYKRAKAVVEGSEGGILPLGTGQGRGLBEAAKKTGHGYVSTRWI  
PGTLNTEISIGIKORIDNDNPTBALSPNETSQVXKPDLLVVLNPTNRALLLE  
AIKSRVPTAIIIDTDEPSLVTPYIPGNDSLRSVNFLLGLARAGQGLNRLARN  
EK"  
complement(1519..2754)  
/gene="LAG1"

CDS  
complement(1519..2754)  
/gene="LAG1"  
/note="YHL003C; Similar to UOG1, hypothetical protein in  
GDF-1.5' region (mouse and human). Null mutant shows  
reduced mean and maximum life span"  
/codon\_start=1  
/evidence=not experimental  
/product="Lag1p"  
/protein\_id="AAB68429.1"  
/db\_xref="GI:500820"  
/translation="MPSATDQSTIDLVVNAKTRRRSSVGKIDLGTVPGPAAPESA  
ASKKRAKRAKALTGDSKSDILMKWPFSTRENNYHSMITPPIIVCYAFSLG  
NRTESNPLHMFVAISQVGDISTYAKGIDKLSFVFETWIFETLRLMDVIRPFTV  
YLVNTSHRQRMLEOMVAFIYCGVSGPGLYIMHSDMLFKTPMYRTYVITNPE  
LFKIFVYGOAFAFMOACAVLQCEKSKDYKEFHIIYLLIMSSYVHFPMGIL  
AIVTMDVSDPEFLSKTLNLYNSVFTPVFGVLEFPFMIYLRHVNRIRIMSVTEFR  
HEGVYVNFPAQCYKWCWISLPIYVLIQAQVNLVNLFLIRLYLRLINQIGIKDER  
SDSDSDSAENESKEKE"  
3482..4840  
/gene="YHL002W"  
3482..4840  
/gene="YHL002W"  
/note="Similar to several proteins with SH3 domains"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="AAB68427.1"  
/db\_xref="GI:500818"  
/translation="MSSSAIKIRNALIKATDPKLRSDMVOYILDVCDLVKEDPDNG  
EWSLIRKRIEQDANVILRTLSITVLAENCSRLQESISKRTSLVALLESYH  
HITLKAIVTVKQSDSFQDDPSLRMGDYLDTIKKAPFLVQPNPEKRMSTQAD  
NSDDEBLQAKLMSLFYEKQKQLQEQEKSAEVLPOQOQHOQOAPAKITAOV  
VRRVRLYDLTNEPDELSPKGDVIVLVEQVDMWKGALRGMMGFLPNVYPIVE  
PSKEIEKKNKEAIVFSOKTIDQLNLSAASKTNSNEVLQDPHIGMYGVTPL  
RPOVTRMLGKYAKKEDMLSLROYLAASVYQNLMDRAANAHISPPVGPALAGMT  
HANTPVPMPQKQSYQSNYSPIYSNLPIDHPNTSANTTYQYGLDLSVYSQPPCY  
EQ"  
5147..5961  
/gene="RPL14B"  
join(5147..5275,5674..5961)  
/gene="RPL14B"  
/note="YHL001W; identical to hypothetical protein YKL006W  
on chromosome 11"  
/codon\_start=1  
/evidence=not experimental  
/product="Rpl14bp: Probable 60S ribosomal protein L14B"  
/protein\_id="AAB68426.1"  
/db\_xref="GI:500817"  
/translation="MSTDSIVKASNRLVEGRVVLIKKQSGAKLAIVETIDQKY  
LIDGPKAGVPRQAINLQGVVLTPTFLPFGARTATVSKMAAGVCEKMAASSWAK  
IAQERRRALTDFFERFQVMVLRKQRYTKKALAKA"  
complement(6456..6480)  
/note="misc feature comment"  
complement(6481..6565)  
/note="misc feature comment"  
complement(6566..6573)  
/note="misc feature comment"  
6925..8238  
/gene="YHR001W"  
6925..8238  
/gene="YHR001W"  
/note="weakly similar to 60 kD chaperonin (Rickettsia  
tsutsugamushi)"  
/codon\_start=1  
/evidence=not experimental  
/product="Yhr001wp"  
/protein\_id="AAB68425.1"  
/db\_xref="GI:500816"  
/translation="MALNKLKKNIPSLTNSHSSINGIASNANSPGADTDIDEND  
ESGGSILNITISQLKPGCDLSRITLPPILIEKSMLEBITNQLQFPVYLAEHNSKQ  
LQRTVRYVNTLACMHIQPRAYVKRPLAPITGEHTAYWDLPNKQAPYILAEQSHHP  
ESAFVMIIPESNLRVGDVVPKSFLLDNSSANAMEGLTVLQFDLIKANGKPEKRTIS  
QPNVYARGILFGKRIELGDHMYMGKYQVDLEFKTGFISGYDAIEGTLKQYDGK  
EYQISGKMDIMYIKDLREKSKKTVLPFHQHFPLAPKVRPLEEGEYESRRLMKK

```

gene
CDS
        /gene="QCR10"
        /join(8696..8701,8756..8992)
        /gene="QCR10"
        /note="YHR001W-A"
        /codon_start=1
        /product="Qcr10p: 8.5 kDa subunit of the quinol cytochrome
        oxidoreductase complex"
        /protein_id="BAB8434.1"
        /db_xref="GI:2351226"
        /translation="MAILQVSHLSKTLHFGRLSLTAYAPNLMWGASMLGL
        FVPEFGPKRQDTLYKKIPLLGPTLEDDHTPEDEKN"
        9682..10755
        /gene="YHR002W"
        /gene="YHR002W"
        /note="YHR002W"
        /protein_id="BAB8424.1"
        /db_xref="GI:500815"
        /translation="MTRDSPDSNDSTYHINKNTOKTSFDRNSFDYIVRSLAGISG
        SCATLIAPLDRIKILFQTSNPHYTKYGLILGVEAKHIMDVGKGFQGSATL
        LAIEPRAVKFVAEJOIRNTLIPSKFESHWRILVSGSLGCVFTYPLDIVRVL
        ALEFHRAVKGRLIIKKIKYPEASATLIKNDYIINWFCWMCNFGVVPYVIGMI
        GVSPFADHLIDHDKSPPAAPVSYLSEDEDELRVOKKORPLRTWAEILSGLAGM
        ASQRAYPAFETIRRLQVSAISPRTMDHKQISIEIAHIIIFKRGVGRFFVGISIGY
        IKQIPWACSPFVIERMKMNGI"
        /complement(10898..12187)
        /gene="YHR003C"
        /gene="YHR003C"
        /complement(10898..12187)
        /note="Weak similarity to thif (E. coli), moeb (E. coli),
        and ubiquitin activating enzyme (wheat)"
        /codon_start=1
        /evidence=not experimental
        /product="Yhr003cp"
        /protein_id="BAB8430.1"
        /db_xref="GI:500821"
        /translation="MANNTWKLATLALISVFSQTLAKSVWEKYLCAANKNTVSR
        PRQVDHLFREQLARNVAFLEEGEMRKIKEDYIIVAGIEVGSVCTMLIRSGCKIM
        IIDENISIDSLNTHCCAVSIDGPKVQCKEHLSKLAPASEIKAKAKMTKENSND
        LIFADGSPFTIVDCILNLSKVDLLFYAHNKRIDVYSMWVAKSPTRVSTINDISM
        TEPPIRCVRRKLRKRGIALTGISIVFSNEMLDRDDLIISPIDCEHRAINADEAL

```

VERSION	AF061031.2	GI:15724846
KEYWORDS	Mycobacterium smegmatis.	
SOURCE	Mycobacterium smegmatis	
ORGANISM	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Acidimicrobiales; Corynebacterineae; Mycobacteriaceae;	
REFERENCE	1 (bases 1 to 1071)	
AUTHORS	Harth,G. and Horwitz,M.A.	
TITLE	Export of recombinant Mycobacterium tuberculosis superoxide dismutase is dependent upon both information in the protein and mycobacterial export machinery. A model for studying export of leaderless proteins by pathogenic mycobacteria	
JOURNAL	J. Biol. Chem. 274 (7), 4281-4292 (1999)	
MEDLINE	99134360	
PUBMED	9933629	
REFERENCE	2 (bases 1 to 1071)	
AUTHORS	Tullius,M.V., Harth,G. and Horwitz,M.A.	
TITLE	High extracellular levels of Mycobacterium tuberculosis glutamine synthetase and superoxide dismutase in actively growing cultures are due to high expression and extracellular stability rather than to a protein-specific export mechanism	
JOURNAL	Infect. Immun. 69 (10), 6348-6363 (2001)	
MEDLINE	21437642	
PUBMED	11553579	
REFERENCE	3 (bases 1 to 1071)	
AUTHORS	Harth,G. and Horwitz,M.A.	
TITLE	Direct Submission	
JOURNAL	Submitted (23-APR-1998) Medicine, UCLA, 10833 Le Conte Avenue, Los Angeles, CA 90095, USA	
REFERENCE	4 (bases 1 to 1071)	
AUTHORS	Tullius,M.V., Harth,G. and Horwitz,M.A.	
TITLE	Direct Submission	
JOURNAL	Submitted (24-JAN-2001) Medicine, UCLA, 10833 Le Conte Ave., Los Angeles, CA 90095, USA	
REMARK	Nucleotide sequence updated by submitter	
COMMENT	On Sep 24, 2001 this sequence version replaced gi:4321777.	
FEATURES	Location/Qualifiers	
source	1..1071	
CDS	/organism="Mycobacterium smegmatis"	
	/strain="1-2c"	
	/db_xref="taxon:1772"	
	<1..141	
	/note="similar to partial orf adjacent to Mycobacterium fortuitum sodA gene"	
	/codon_start=1	
	/transl_table=11	
	/product="unknown"	
	/protein_id="AAU06643.1"	
	/db_xref="GI:15724848"	
	/translation="GSGSALPQPEWHTSHGLQTRHGVAFRLACVIAALLGLVAFLITWF"	
gene	282..905	
	/gene="sodA"	
CDS	282..905	
	/gene="sodA"	
	/EC_number="1.15.1.1"	
	/codon_start=1	
	/transl_table=11	
	/product="superoxide dismutase"	
	/protein_id="ADI5825.2"	
	/db_xref="GI:15724847"	
	/translation="MAEYTLPLDLVDYGALEPHTSGOINELHSHKHAATYYKGVDNALAKLEPRANGDHAIPLNEKILAFHLGGHNHSIWNKSLNSNGDKPTGEIAALIDPOCFESDFEQAPTAAANGLGSGWAIVAYDLSLGGLLTFQLYDQANVPGLTPILOVDVMHEADFTLYQIKNVKADVKAFAWVVNWDVDQNREPAATSKTSGLIFG"	
	964..996	
BASE COUNT	205 a 349 c 315 g 202 t	
ORIGIN		

Percent Similarity: 89.29%  
 Best Local Similarity: 57.14%  
 Query Match: 68.28%  
 DB: 1  
 Gaps: 0

US-09-987-190-2 (1-30) x AF061031 (1-1071)

Qy 1 LysTYSerLeuProGluLeuAspTYrGluPhseSerAlaThrGluProTYrIleSerGly 20  
 Db 268 GAATACACCTCCCGAGCTGATTAACGATGAGCCCTCGAGCCCATATCTCCGCT 347

Qy 21 GlnIleAsnGluIle\*\*TYrThr 28  
 Db 348 CAGATCAACGAGCTCCACACAGC 371

## RESULT 13

LOCUS MFSODA 1478 bp DNA linear BCT 21-MAY-1996  
 DEFINITION M.fortuitum sod gene.  
 ACCESSION X70914  
 VERSION X70914.1 GI:950078  
 KEYWORDS sod gene; superoxide dismutase.  
 SOURCE Mycobacterium fortuitum.  
 ORGANISM Mycobacterium fortuitum  
 Bacteria; Firmicutes; Actinobacteriae; Actinobacteridae;  
 Actinomycetales; Corynebacterineae; Mycobacteriaceae;  
 Mycobacterium.

REFERENCE 1 (bases 1 to 1478)  
 AUTHORS Menendez,M.C., Domenech,P., Prieto,J. and Garcia,M.J.  
 TITLE Cloning and expression of the Mycobacterium fortuitum superoxide  
 JOURNAL dismutase gene  
 MEDLINE FEMS Microbiol. Lett. 134 (2-3), 273-278 (1995)  
 PUBMED 8586272

REFERENCE 2 (bases 1 to 1478)  
 AUTHORS Menendez,C.  
 TITLE Direct Submission  
 JOURNAL Submitted (15-FEB-1993) C. Menendez, School of Medicine, Autonoma  
 University of Madrid, Dept. of M. Preventive School of Medicine,  
 ST/Arzobispo Morcillo, 4, 28029 Madrid, SPAIN  
 revised by [3] MAT  
 3 (bases 1 to 1478)  
 AUTHORS Garcia,M.  
 TITLE Direct Submission  
 JOURNAL Submitted (17-AUG-1995) M.Garcia, School of Medicine, Autonoma  
 University of Madrid, Dept. of M. Preventive School of Medicine,  
 ST/Arzobispo Morcillo, 4, 28029 Madrid, SPAIN

## FEATURES

## source

1..1478  
 /organism="Mycobacterium fortuitum"  
 /strain="ATCC 6841"  
 /db\_xref="taxon:1766"  
 /clone="pmFV1"  
 /clone\_1id="E.coli TG1"  
 243..250  
 /gene="Sod"  
 257..880  
 /gene="Sod"  
 257..880  
 /gene="Sod"  
 /codon\_start=1  
 /transl\_table=11  
 /product="superoxide dismutase"  
 /protein\_id="CAA50266.1"  
 /db\_xref="GI:950078"  
 /db\_xref="SWISS-PROT:Q59519"  
 /translation="MAEYTLPLDLDYDALPEPHISGQINELHSHKHAAYVKGVNDAY  
 AKLDEARANGDHAAIPLNEKNLAFLGSHVNSIWMKNLSPNGGDKPTGLADIDET  
 FGSPDKFOAFTAAANGLOGSGMAVILGSDYDLGNKLLTFOLDDOQANVPLGIIPLLQVD  
 MMEHAFYLOYNVADYVKAFMNVVNMEDVQNRVAAATSKNGLIIFG"

BASE COUNT 260 a 486 c 452 g 280 t

Alignment Scores:

Pred. No.: 4,61e-08  
 Score: 99.00  
 Percent Similarity: 89.29%  
 Best Local Similarity: 57.14%  
 Query Match: 68.28%  
 DB: 1  
 Gaps: 0

US-09-987-190-2 (1-30) x MFSODA (1-1478)

Qy 1 LysTYSerLeuProGluLeuAspTYrGluPhseSerAlaThrGluProTYrIleSerGly 20  
 Db 263 GAATACACTTTCGCGAGCTGATTAACGATGAGCCTCGAGCCCATATCTCCGCG 322

Qy 21 GlnIleAsnGluIle\*\*TYrThr 28  
 Db 323 CAGATCAACGAGCTCCACACAGC 346

## RESULT 14

LOCUS MLEPSOD 649 bp DNA linear BCT 10-FEB-1999  
 DEFINITION Mycobacterium leprae DNA for manganese superoxide dismutase (EC  
 1.15.1.1).  
 ACCESSION X16453  
 VERSION X16453.1 GI:44405  
 KEYWORDS manganese superoxide dismutase; metalloenzyme; superoxide  
 dismutase.  
 SOURCE Mycobacterium leprae.  
 ORGANISM Mycobacterium leprae  
 Bacteria; Firmicutes; Actinobacteriae; Actinobacteridae;  
 Actinomycetales; Corynebacterineae; Mycobacteriaceae;  
 Mycobacterium.

REFERENCE 1 (bases 1 to 649)  
 AUTHORS Thangaraj,H.S.  
 TITLE Direct Submission  
 JOURNAL Submitted (01-SEP-1989) Thangaraj H.S., National Institute for  
 Medical Research, Laboratory of Leprosy and Mycobacterial Research,  
 Mill Hill, London NW7 1AA, UK  
 2 (bases 1 to 649)  
 AUTHORS Thangaraj,H.S., Lamb,F.I., Davis,E.O. and Colston,M.J.  
 TITLE Nucleotide and deduced amino acid sequence of Mycobacterium leprae  
 JOURNAL manganese superoxide dismutase  
 MEDLINE Nucleic Acids Res. 17 (20), 8378 (1989)  
 PUBMED 2682526

## FEATURES

## source

1..649  
 /organism="Mycobacterium leprae"  
 /db\_xref="taxon:1769"  
 /clone="Y3164 and pH77"  
 /clone\_1id="lambda gcl1 and pHCT9 cosmid"  
 1..4  
 /note="poc. ribosome binding site"  
 14..637  
 /note="superoxide dismutase (AA 1-207)"  
 /note=""  
 /codon\_start=1  
 /transl\_table=11  
 /protein\_id="CAA34472.1"  
 /db\_xref="GI:581343"  
 /db\_xref="SWISS-PROT:P13367"  
 /translation="MAEYTLPLDLDYDALPEPHISGQINELHSHKHAAYVKGVNDAL  
 AKLDEARAKDHSNIFLNEKNLAFLGSHVNSIWMKNLSPNGGDKPTGLADIDET  
 FGSPDKRAQFSAANGLQSGMAVILGSDYDLGNKLLTFOLDDOQANVPLGIIPLLQVD  
 MMEHAFYLOYNVADYVKAFMNVVNMEDVQNRVAAATSKNGLIIFD"

BASE COUNT 164 a 180 c 167 g 138 t

Alignment Scores:

Pred. No.: 2.53e-08  
 Score: 98.00  
 Percent Similarity: 92.86%  
 Best Local Similarity: 57.14%  
 Query Match: 67.59%  
 DB: 1  
 Gaps: 0





```

/codon_start=1
/transl_table=1
/product="putative DNA replication and SOS induction
protein"
/protein_id="CAC29511.1"
/db_xref="GI:13092415"
/translation="MYVHFGRLRDFRSMHDVLELNPGRVTFPGPNNGKTNLLIEMW
YSTLSHRVGTDIPLIRAGTIRAIIVSTIVNERECIDIEIAGRANRRLRSIV
RGMREVGVLRAVLEAPEDLALVGDGPANRRVYDLATVOPVIAARADYKVLRO
RALLKSLAARVSRDOGVLDLVDWDRRLRHEAGELMAAIDILVNOLEPVEKAYOL
LAPGSRSTISIRASLDIGTACGVSSRRALLQDLADLAGSTRNVELERICVGRH
RDELELRIGDQPAKGFASHGWSLALRLIAAYELLADGNEVLLDLDVFAELDAA
RCRALATVAESAEQVLTYSAAQEDI PVGMADKMTVDLRDSGRVSVYP"
3287..3415
/gene="recF"
/ncore="Pfam match to entry PF00470 RecF, RecF protein,
score 18.90, E-value 0.00024"
3368..3391
/gene="recF"
/ncore="PS00017 ATP/GTP-binding site motif A (P-loop)"
3614..3691
/gene="recF"
/ncore="PS00617 RecF protein signature 1"
4100..4423
/gene="recF"
/ncore="Pfam match to entry PF00470 RecF, RecF protein,
score 5.20, E-value 1.5"
4244..4297
/gene="recF"
/ncore="PS00618 RecF protein signature 2"
4435..5004
/gene="ML0004"
4435..5004
/gene="ML0004"
/ncore="Similar to hypothetical proteins from mycobacteria
e.g. M. tuberculosis Rv0004, hypothetical protein,
TR:P71573 (EMBL:AL123456) (187 aa); Fasta score E(): 0,
77.2% identity in 167 aa overlap. Previously sequenced as
TR:050181 (EMBL:Z70722) (199 aa); Fasta score E(): 0,
99.5% identity in 187 aa overlap."
/codon_start=1
/transl_table=1
/product="conserved hypothetical protein"
/protein_id="CAC29512.1"
/db_xref="GI:13092416"
/translation="MIESNESYSGDITIEPIGLTSGPDIVRPALEPARRAACAQGD
AGRGHVPPVPRFTDRRRNMSGPEPPVDRDPOPGAKVAHDIAKRGMSAOYAEGRVRS
QWASWVGQIADHAFVGLNGVLSVTNESTAMATQIKIWOQLAKTIAAIVGNGVVT
SLKITGPTAPSRMRGPMHIAGRGPRDTG"
5211..5216
/ncore="possible RBS"
5229..7265
/gene="gyrB"
/ncore="gyrB"
/ncore="ML0005"
5229..7265
/gene="gyrB"
/ncore="gyrB"
/BC_number="5.99.1.3"
/ncore="Similar to M. tuberculosis gyrB, DNA gyrase subunit
B, SW:GYRB_MYCTU (P41514) (686 aa); Fasta score E(): 0,
87.5% identity in 679 aa overlap and to Mycobacterium
smegmatis gyrB, SW:GYRB_MYCSM (P48355) (675 aa); Fasta
score E(): 0, 85.0% identity in 679 aa overlap. Previously
sequenced as SW:GYRB_MYCLE (Q59533) (697 aa); Fasta score
E(): 0, 100.0% identity in 678 aa overlap. Contains Pfam
match to entry PF00986 DNA_gyraseB_C, DNA gyrase B
subunit, carboxyl terminus. Contains Pfam match to entry
PF00204 DNA_topoisomII, DNA topoisomerase II (N-terminal

```

```

Query Match: 67.59% Indels: 0
DB: 1 Gaps: 0
US-09-987-190-2 (1-30) x MLEPRTNI (1-344050)

Qy 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrTlleSerGly 20
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 93264 GAATACACCCCTCCGCAATTTGAGCTGCGATACGACGCTGGAACACATATCTTGGT 93205
Qy 21 GlnIleAngIuIle**TyrThr 28
Db 93204 GAGTCAACGAGATCCACACACC 93181

RESULT 16
MTSOD MTSOD 790 bp DNA linear BCT 03-MAY-1995
LOCUS M.tuberculosis SOD gene.
DEFINITION X52861.1 GI:794079
ACCESSION X52861.1
VERSION sod gene; superoxide dismutase.
KEYWORDS Mycobacterium tuberculosis.
SOURCE Mycobacterium tuberculosis
ORGANISM Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium tuberculosis complex.
1 (bases 1 to 790)
Zhang, Y.
REFERENCE
AUTHORS Submitted (23-APR-1990) Zhang Y., MRC TB & Related Infections Unit,
TITLE Cyclotron Bldg, Hammersmith Hospital, London W12 0HS, U K
JOURNAL 2 (bases 1 to 790)
AUTHORS Zhang, Y., Lathigra, R., Garbe, T., Catty, D. and Young, D.
TITLE Genetic analysis of superoxide dismutase, the 23 kilodalton antigen
of Mycobacterium tuberculosis
Mol. Microbiol. 5 (2), 381-391 (1991)
JOURNAL MEDLINE 91251768
PUBMED 1904126
COMMENT On May 4, 1995 this sequence version replaced gi:44678.
Data kindly reviewed (02-JUL-1990) by Zhang Y.
Related sequence: S36714.

FEATURES
Source
1..790
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
/clone="YA3P-9"
/clone_lib="lambda gt11"
20..25
/ncore="putative"
99..105
112..735
/gene="SOD"
112..735
/gene="SOD"
/BC_number="1.15.1.1"
/BC_start=1
/codon_start=1
/transl_table=1
/product="superoxide dismutase"
/protein_id="CAA57042.1"
/db_xref="GI:581379"
/db_xref="SWISS-PROT:P17670"
/translation="MAEYTPDLDWDVYDGLPEPHISGOINELHSHKHATYVKGANDAV
AKLEERAKEDHSAIILNEKNAFNLAGVNHNTIMKULSPNGSKRTGELAAIADA
FGSDKRAQFHAATTVQSSGNAALGMDTLGNKLLIFQYDHTQTNPLGIVPLLLD
MWEHAFYLDQKNKVKVDPAKFMVNVNADVDQSRVAATSTQKGLFG"

BASE COUNT 182 a 245 c 226 g 137 t
ORIGIN
Alignment Scores:
Pred. No.: 1.72e-07 Length: 790
Score: 94.00 Matches: 15
Percent Similarity: 89.29% Conservative: 10

```

Best Local Similarity: 53.57% Mismatches: 3  
Query Match: 64.83% Indels: 0  
DB: 1 Gaps: 0

US-09-987-190-2 (1-30) x MTSOD (1-790)

Qy 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20  
Db 118 GAATACACCTTGCAGACTGGAGCTACGAGACCTGAACCCGACATCTCGGGT 177  
Qy 21 GlnIleAsnGluIle\*\*\*TyrThr 28  
Db 178 CAGATCAACGAGCTTCACACAGC 201

RESULT 17  
LOCUS AF077406 943 bp DNA linear BCT 26-JUL-1998  
DEFINITION Mycobacterium bovis BCG superoxide dismutase (SOD) gene, complete cds.  
ACCESSION AF077406  
VERSION AF077406.1 GI:3342171  
KEYWORDS Mycobacterium bovis BCG.  
SOURCE Mycobacterium bovis BCG.  
ORGANISM Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex.  
REFERENCE 1 (bases 1 to 943)  
AUTHORS Kimble, E., Sanderson, R.J. and Gill, R.E.  
TITLE Superoxide dismutase of M. bovis BCG  
JOURNAL Unpublished  
PUBMED 2 (bases 1 to 943)  
REFERENCE Kimble, E., Sanderson, R.J. and Gill, R.E.  
TITLE Direct Submission  
AUTHORS Submitted (10-JUL-1998) Microbiology, University of Colorado Health Sciences Center, 4200 E. 9th Ave., Denver, CO 80220-3706, USA  
JOURNAL Location/Qualifiers  
FEATURES  
source 1..943  
/organism="Mycobacterium bovis BCG"  
/strain="Pasteur"  
/db\_xref="taxon:33892"  
gene 307..943  
/gene="SOD"  
RBS 307..314  
/gene="SOD"  
/note="putative"  
CDS 320..943  
/gene="SOD"  
/note="similar to Mycobacterium tuberculosis superoxide dismutase encoded by the sequence presented in GenBank  
Accession Number X52861"  
/codon\_start=1  
/transl\_table=1  
/product="superoxide dismutase"  
/protein\_id="AAC27527.1"  
/db\_xref="GI:3342172"  
/translation="MAEYTLPLDMDVYGALEPHISGOINELHSHKHATVYKANDAV AKLEBARAKEDSHAILINENKLAFNLAGHNHTIWMKNISPNCGDKPTGELAAIADA FGSFDRAPRAPHAAATTVOSSGNALGMDTLGNKLLIFQYDHTQNFPLGIVPLLLD MMEHAFYLYQYKNKVDPAKAFMNVNADVQSRVAATSTQKGLIFG"

BASE COUNT 211 a 302 c 266 g 164 t

ORIGIN

Alignment Scores:  
Pred. No.: 2.14e-07 Length: 943  
Score: 94.00 Matches: 15  
Percent Similarity: 89.29% Conservative: 10  
Best Local Similarity: 53.57% Mismatches: 3  
Query Match: 64.83% Indels: 0  
DB: 1 Gaps: 0

US-09-987-190-2 (1-30) x AF077406 (1-943)

Qy 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20  
Db 326 GAATACACCTTGCAGACTGGAGCTACGAGACCTGAACCCGACATCTCGGGT 385  
Qy 21 GlnIleAsnGluIle\*\*\*TyrThr 28  
Db 386 CAGATCAACGAGCTTCACACAGC 409

RESULT 18  
LOCUS AF061030 1321 bp DNA linear BCT 03-MAR-1999  
DEFINITION Mycobacterium tuberculosis superoxide dismutase (sodA) gene, complete cds.  
ACCESSION AF061030  
VERSION AF061030.1 GI:4321775  
KEYWORDS Mycobacterium tuberculosis.  
SOURCE Mycobacterium tuberculosis.  
ORGANISM Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex.  
REFERENCE 1 (bases 1 to 1321)  
AUTHORS Hart, G. and Horwitz, M.A.  
TITLE Export of recombinant Mycobacterium tuberculosis superoxide dismutase is dependent upon both information in the protein and mycobacterial export machinery. A model for studying export of leaderless proteins by pathogenic mycobacteria  
JOURNAL J Biol. Chem. 274 (7), 4281-4292 (1999)  
PUBMED 99134360  
MEDLINE 9933629  
REFERENCE 2 (bases 1 to 1321)  
AUTHORS Hart, G. and Horwitz, M.A.  
TITLE Direct Submission  
AUTHORS Submitted (23-APR-1998) Medicine, UCLA, 10833 Le Conte Avenue, Los Angeles, CA 90095, USA  
JOURNAL Location/Qualifiers  
FEATURES  
source 1..1321  
/organism="Mycobacterium tuberculosis"  
/strain="Erdrman"  
/db\_xref="taxon:1773"  
gene 542..1165  
/gene="soda"  
CDS 542..1165  
/gene="soda"  
/EC\_number="1.15.1.1"  
/codon\_start=1  
/transl\_table=1  
/product="superoxide dismutase"  
/protein\_id="AAD15824.1"  
/db\_xref="GI:4321776"  
/translation="MAEYTLPLDMDVYGALEPHISGOINELHSHKHATVYKANDAV AKLEBARAKEDSHAILINENKLAFNLAGHNHTIWMKNISPNCGDKPTGELAAIADA FGSFDRAPRAPHAAATTVOSSGNALGMDTLGNKLLIFQYDHTQNFPLGIVPLLLD MMEHAFYLYQYKNKVDPAKAFMNVNADVQSRVAATSTQKGLIFG"

BASE COUNT 266 a 429 c 382 g 244 t

ORIGIN

Alignment Scores:  
Pred. No.: 3.25e-07 Length: 1321  
Score: 94.00 Matches: 15  
Percent Similarity: 89.29% Conservative: 10  
Best Local Similarity: 53.57% Mismatches: 3  
Query Match: 64.83% Indels: 0  
DB: 1 Gaps: 0

US-09-987-190-2 (1-30) x AF061030 (1-1321)

Qy 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20  
Db 548 GAATACACCTTGCAGACTGGAGCTACGAGACCTGAACCCGACATCTCGGGT 607  
Qy 21 GlnIleAsnGluIle\*\*\*TyrThr 28  
Db 1 Gaps: 0

Db	608	CAGATCAACGAGCTTCACCACGC	631
RESULT 19	AE007188	14282 bp	DNA linear BCT 27-APR-2001
LOCUS	AE007188		
DEFINITION	Mycobacterium tuberculosis CDC1551, complete genome.		
ACCESSION	AE007188	AE000516	
VERSION	AE007188.1	GI:13883836	
KEYWORDS			
SOURCE			
ORGANISM	Mycobacterium tuberculosis CDC1551.		
REFERENCE	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex.		
AUTHORS	1 (bases 1 to 14282) Fleischmann, R.D., Alland, D., Eisen, J.A., Carpenter, L., White, O., Peterson, J., Deboy, R., Dodson, R., Gwinn, M., Hatt, D., Hickey, E., Kolonay, J.F., Nelson, W.C., Umayam, L.A., Ermolaeva, M., Salzberg, S.L., Delcher, A., Utterback, T., Weidman, J., Khouri, H., Gill, J., Mikula, A. and Bishai, W.		
TITLE	Whole genome comparison of Mycobacterium tuberculosis clinical and laboratory strains		
JOURNAL	unpublished		
REFERENCE	2 (bases 1 to 14282) Fleischmann, R.D., Alland, D., Eisen, J.A., Carpenter, L., White, O., Peterson, J., Deboy, R., Dodson, R., Gwinn, M., Hatt, D., Hickey, E., Kolonay, J.F., Nelson, W.C., Umayam, L.A., Ermolaeva, M., Salzberg, S.L., Delcher, A., Utterback, T., Weidman, J., Khouri, H., Gill, J., Mikula, A. and Bishai, W.		
AUTHORS	Direct Submission Submitted (25-APR-2001) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA		
FEATURES	Location/Qualifiers		
source	1..14282		
gene	/organism="Mycobacterium tuberculosis CDC1551"		
CDs	/strain="CDC1551"		
gene	/db_xref="taxon:83331"		
CDs	/note="clinical strain"		
gene	130..675		
CDs	/gene="MT3949"		
gene	130..675		
CDs	/gene="MT3949"		
gene	/note="similar to GB:AE000666; identified by sequence similarity; putative"		
CDs	/codon_start=1		
gene	/transl_table=1		
CDs	/product="ferritin family protein"		
gene	/protein_id="AAK4316.1"		
CDs	/db_xref="GI:13883837"		
gene	/translation="MTYEGKPKFKALMQEIHNEFTAAQOYVAIAYFPSEDLPOIAKHQFSQAVERRNHAMLVQHLDDLDLVEIIGVDVTVNGDGRREALALDDERTYTDQGRFLAVARDEGDFEGEFGMQLFEOQLBEVALMATTVRVADRAGANFLELENFVAREVDAPASGAPPAAGRL"		
CDs	complement (690..1520)		
gene	/gene="MT3950"		
CDs	complement (690..1520)		
gene	/note="similar to SP:P37965.GB:Z26522.PID:403373GB:AE00126; identified by sequence similarity; putative"		
CDs	/codon_start=1		
gene	/transl_table=1		
CDs	/product="glycerophosphoryl diester phosphodiesterase, putative"		
gene	/protein_id="AAK4317.1"		
CDs	/db_xref="GI:13883838"		
gene	/translation="MDMTMADEVLVAGHPVVAHRCASARPEHTLAAYDLALKEGADGQVECVRLTRDGHILVCVMDRLDRSTAGALVSTMTLQRLRELEGAMDSWRPDGSHGDSLTLLDALVSLVDHMRPVKIFVETGHPVRSYSLVNTKLLALHRCGIAAPASABRSRAVMSAAVWRIRARPAFLPVYILGKTPRYTSSAAAVAGATNGPBLPALKEYEPQVDRSAAQGAAYTCANVDEYEDIDFCREVGAMVIGTHFGRTKAMLEDGRANGTTR"		
CDs	complement (1520..2548)		

```

CDS      /gene="WT3951"  

          complement(1520..2548)  

          /gene="WT3951"  

          /note="identified by Glimmer2; putative"  

          /codon_start=1  

          /transl_table=11  

          /product="hypothetical protein"  

          /protein_id="AAK48318.1"  

          /db_xref="GI:13883839"  

          /translation="MIQISQCGTGMNVRERORWVCBRCGMLLAPLADMPAPARWET  

          PAVDPAPASTRRTPRLPGLPGFWRIVAGCAAPPPHGRRLRGPTRYAGI PRWGLT  

          HVDVDPAPSAKAGSPAAARTLLVLSLVFSIAVFPVVRVYLIVINNTLLVGL  

          ASASVWGLVVSLSAIAAGCTIVLYIRMTVARAAAFMQGAPERSASEIWAACIL  

          PMVNTLVVIEALVEDRYTLRRLPYVMWIVTSNAISGFAPATSYVDQIGI  

          ANNTIMWMLALCAAAVAARVBEFGEQKPEVRPAHRVWVNTDGRSAPASSVAVE  

          LDGEPPAA"  

gene     2776..2880  

CDS      /gene="WT3952"  

          2776..2880  

          /gene="WT3952"  

          /note="identified by Glimmer2; putative"  

          /codon_start=1  

          /transl_table=11  

          /product="hypothetical protein"  

          /protein_id="AAK48319.1"  

          /db_xref="GI:13883840"  

          /translation="WDNSARQVLIDSLAGTRAKVIALSIHCPGCT"  

          3087..3236  

gene     /gene="WT3953"  

          3087..3236  

          /gene="WT3953"  

          /note="identified by Glimmer2; putative"  

          /codon_start=1  

          /transl_table=11  

          /product="hypothetical protein"  

          /protein_id="AAK48320.1"  

          /db_xref="GI:13883841"  

          /translation="WATHSYGPSSTTPNGSGSPENHAYGIALRALIALLAIIIA  

          VIALV"  

gene     complement(3637..3849)  

          /gene="WT3954"  

          complement(3637..3849)  

          /gene="WT3954"  

          /note="identified by Glimmer2; putative"  

          /codon_start=1  

          /transl_table=11  

          /product="hypothetical protein"  

          /protein_id="AAK48321.1"  

          /db_xref="GI:13883842"  

          /translation="MAIRVLDLGRKDSGRPIQDSPGLVDAGIHGAVTAAELAK  

          EFCLGACVAGIRAIPIGSLDRAAQVQ"  

gene     complement(4199..4801)  

          /gene="WT3954.1"  

          complement(4199..4801)  

          /gene="WT3954.1"  

          /note="identified by Glimmer2; putative"  

          /codon_start=1  

          /transl_table=11  

          /product="hypothetical protein"  

          /protein_id="AAK48322.1"  

          /db_xref="GI:13883843"  

          /translation="WVTGRDRRVDSDORPPAPSWILGSHGLALCMEVLPIYVNVVL  

          MESSIIPPTPTVRELVAIGESPITALPAGLPNHRPQGGSSRAAPVLVDSMWH  

          PSTRCCALPQGHQFPKQGFRRQSSRSLSAOTLISKRRLAGLPBPVYLTTP  

          SKRHAKYKPEKYAHSGRLEDPQVQCLTLTFARDRQVP"  

gene     4727..5218  

          /gene="WT3955"  

          4727..5218  

          /gene="WT3955"  

          /note="similar to GP.1841455; identified by sequence  

          similarity; putative"  

          /codon_start=1  

          /transl_table=11

```

```

/product="IS1608", transposase"
/protein_id="AAK48323.1"
/db_xref="GI:13883844"
/translation="MTAENPGSRRTLVGIDAITACGHLAIRDDVGARSIRSEVEPT
LAGRLTLDKSGVDIDATVEPTSMWMLPLTIVENAGDTMAGARHCHRLGALV
GKSKSDVIDAEVLTRASEVDFLTPLTPTAQLALRRSVIRRAQAVIDARRSWRLMS
LAR"
complement(5255..5797)
/gene="MTJ3957"
complement(5255..5797)
/gene="MTJ3957"
/feature="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK48324.1"
/db_xref="GI:13883845"
/translation="MGLSKTIPRCRFLPLVGRGRSSSLRCGDTYTPGSPPTMGK
EPGHEITASSDRVCNRIPCRQMSACRFVRLHGVAVVWVGVDVGVQFPR
AAGGDRVAVGVAQLVGAAGDRHPRGAGLQAGDGLGVAMPVGHQPVVERQVAG
RCAGQQADVGPPTPQRCGR"
5778..5993
/gene="MTJ3958"
5778..5993
/gene="MTJ3958"
/feature="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK48325.1"
/db_xref="GI:13883846"
/translation="MVPDKPTVCLSVSHFORLFRVAQNHMPVEIRRDYTHQHLDH
RDSGRRLTSSFAFPADATFORHSS"
complement(6207..6386)
/gene="MTJ3959"
complement(6207..6386)
/gene="MTJ3959"
/feature="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK48326.1"
/db_xref="GI:13883847"
/translation="MMRRSMGWGKKKSTAGLQIAGTANETKEVLEBAVHRESPVIRP
DVVVGIPAVDRRPRQ"
6657..7280
/gene="MTJ3960"
6657..7280
/gene="MTJ3960"
/feature="similar to SP:P47201 PID:555746 PID:620090
PID:667039 SP:P53647, identified by sequence similarity;
putative"
/codon_start=1
/transl_table=11
/product="superoxide dismutase"
/protein_id="AAK48327.1"
/db_xref="GI:13883848"
/translation="MAEYTLPDLDMDVGALEPHISQINLEHLSKHATYVKANDAV

```

## Alignment Scores:

```

Pred. No.: 6.22e-06 Length: 14282
Score: 94.00 Matches: 15
Percent Similarity: 89.29% Conserves: 10
Best Local Similarity: 53.57% Mismatches: 3
Query Match: 64.83% Indels: 0
DB: 1 Gaps: 0

```

US-09-987-190-2 (1-30) x AE007188 (1-14282)

```

Cy 1 LysTyrSerLeuProGluLeuAspTyrGluPheserAlaThrGluProTyrIleSerGly 20
Db 6663 GAATACACCTTGCCGACCTGAGACTACGAGACCTGGAACCGCACATCTCGGGT 6722

```

```

Cy 21 GlnIleAsnGluIle***TyrThr 28
Db 6723 CAGATCAACGAGCTTACACACG 6746
RESULT 20
MTCY1A6
LOCUS 37751 bp DNA linear BCT 03-AUG-2001
DEFINITION Mycobacterium tuberculosis H37Rv complete genome; segment 159/162.
ACCESSION 283864 AL123456
VERSION 283864.1 GI:3261687
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
1 (bases 1 to 37751)
Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C.,
Harris, D., Gordon, S.V., Eigmeier, K., Gas, S., Barry III, C.B.,
Tekala, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T.,
Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S.,
Hamlin, N., Holroyd, S., Hornby, T., Jagels, K., Krogh, A., McLean, J.,
Moule, S., Murphy, L., Oliver, S., Osborne, J., Quail, M.A.,
Rajandream, M.A., Rogers, J., Rutter, S., Seeger, K., Skellon, S.,
Squares, S., Squares, R., Sulston, J.E., Taylor, K., Whitehead, S. and
Barrall, B.G.
Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence
Nature 393 (6685), 537-544 (1998)
9634230
2 (bases 1 to 37751)
Parkhill, J.
Direct Submission
Submitted (11-JUN-1998) Submitted on behalf of the Mycobacterium
tuberculosis sequencing and mapping teams, Sanger Centre, Wellcome
Trust Genome Campus, Hinxton, Cambridge CB10 1SA Unité de Génétique
Moléculaire Bactérienne, Institut Pasteur, 28 rue du Docteur Roux,
75724 Paris Cedex 15, France E-mail: parkhill@sanger.ac.uk
On Jun 27, 1998 this sequence version replaced gi:1781088.
Notes:
Details of M. tuberculosis sequencing at the Sanger Centre are
available on the World Wide Web.
URL: http://www.sanger.ac.uk/Projects/M_tuberculosis/ CDS have
been renumbered from the original cosmid submissions but the old
gene designations are in brackets after the new gene numbers.
Gene prediction was based on a Hidden Markov Model of TB genes
implemented in TBparse (Krogh) supplemented with visual inspection
of positional base preference in codons, especially where there is
an increase in the observed/expected third position G + C.
CAUTION: In some cases we may not have predicted the correct
initiation codon. Where possible we choose an initiation codon
(atg, gtc, or tgc) which is preceded by an upstream ribosome
binding site sequence (optimally 5-13bp before the initiation
codon). If this cannot be identified we choose the most upstream
initiation codon.
FEATURES
source
1..37751
location/Qualifiers
1..37751
/organism="Mycobacterium tuberculosis H37Rv"
/strain="H37Rv"
/db_xref="taxon:83332"
<1..37447
/organism="Mycobacterium tuberculosis H37Rv"
/strain="H37Rv"
/db_xref="taxon:83332"
/clone="Y1A6"
complement(48..677)
/gene="RV3830c"
complement(48..677)
/gene="RV3830c"
complement(48..677)
/gene="RV3830c"
/feature="transcriptional regulator, similar to TCMR-STGGA P39885
tetracenomycin c transcriptional repressor (226 aa), faasta

```

```

scores, opt: 255, E(): 1.9e-11, (33.7% identity in 202aa
overlap). Contains possiblehelix-turn-helix motif from aa
34-55 (+4.97 SD)
/codon_start=1
/translate=1
/product="hypothetical protein Rv3830c"
/protein_id="CAB06211.1"
/db_xref="GI:1781127"
/db_xref="SPTREMBL:P96248"
/translation="MVRPPQTARSRTTEALRQALVRFLLAQGVETSAGQIAEDACV
SLRTFYFRFSKHDLFPADYDAGLHWFRPAUDARPADSITDSVOALFSPYVDAY
TKIASLRGGELEPSRIYVRHREVEADPADALQAO LRNRNCIDIAQAPDARLHIAVTAR
VAAVFGAMEAMLGSDRSGLARVCHVALESIRVGISDWTTLTVSS"
749..1231
/gene="Rv3831"
749..1231
/gene="Rv3831"
/note="Rv3831, (MTCY01A6.38c), len: 160. Unknown"
/codon_start=1
/translate=1
/product="hypothetical protein Rv3831"
/protein_id="CAB06210.1"
/db_xref="GI:1781126"
/db_xref="SPTREMBL:P96247"
/translation="MVSLLVHAALGVVIVGIVSNPKVFTPRPAGSWSFLPECVYV
VGIASIALGMYENIRFVQVYVAGALPLMGSGMAEYRLMTNPPAASSAGODYTIAN
VILPLFSTDTGVRGLRPLVYFSSLFTSFAPAFAYFATIERQHRHRSRATVGA
"
complement(1228..1803)
/gene="Rv3832c"
complement(1228..1803)
/gene="Rv3832c"
/note="Rv3832c, (MTCY01A6.37), len: 191. Unknown."
/codon_start=1
/translate=1
/product="hypothetical protein Rv3832c"
/protein_id="CAB06209.1"
/db_xref="GI:1781125"
/db_xref="SPTREMBL:P96246"
/translation="MAMMLLHRHSSAGMKAVANOLLPMALOHVELGPTLTIGRC
YGATLQALGLTASITLAVVDNMSVNERNRYGQARLITRGDGTGTGPDHFTSVVC
FTMLHVAASQLODOLFAEAYRVLPQGVGSGVPSLPRLIHIADTYPIAPADL
PGRRAVGFDTIHVDVAGARLWRATKEVAA"
1859..2650
/gene="Rv3833"
1859..2650
/gene="Rv3833"
/note="Rv3833, (MTCY01A6.36c), len: 263. Possible
transcriptional regulator, similar to e.g. YFIF_BACSU
P54722 hypothetical transcriptional regulator (314 aa),
fasta scores, opt: 171, E(): 8.5e-05, (30.2% identity in
126 aa overlap). Contains probable helix-turn-helix motif
from aa 165-186."
/codon_start=1
/translate=1
/product="hypothetical protein Rv3833"
/protein_id="CAB06208.1"
/db_xref="GI:1781124"
/db_xref="SPTREMBL:P96245"
/translation="MSENSHRLATITSLTPGARIERHRHSHQIVYPSAGAVSTT
HAGTWTFVNRRAIWI PACMWHQKHFGHTQHGVALDQRGRGAPPTVAVPLMK
ELVIAQSQADRTDDEHRRMLVLODOLPTTSIREPLVPSPTRRRLHACALADNL
TOPLLTIOIGRIGVSORTLSRLPSDELGMFPQWRQTLRQHALVLAERHDVTSVA
SEGCWATPSAIDTYRQAFGHTPGQAAKPMATRLTRLRARDKR"
complement(2647..3906)
/gene="sers"
complement(2647..3906)
/gene="sers"
/note="Rv3834c, (MTCY01A6.35), len: 419. Seryl-trna
synthetase, similar eg to SIS_BACU P37464 (425 aa), fasta
scores, opt: 1015, E(): 0, (39.3% identity in 425 aa
overlap); contains P500179 Aminoacyl-transfer RNA
synthetases class-II signature 1"

```

```

/codon_start=1
/translate=1
/product="sers"
/protein_id="CAB06207.1"
/db_xref="GI:1781123"
/db_xref="SWISS-PROT:P96244"
/translation="MIDLKLRNPDAVRNRSQLSRGEPALVDALLTADARRAVIST
ADSLRAEQKAASKVGASPEERPLRLRAKELEQVKAADVEVAEAAFTAAHLAI
SNVIQGVPAAGEDDYAVLDVGEPSYLENRKDLGELGSLGIDMORGAKVSGRPY
FLRGALLQLGLLQALKLVANGFVPTPLVAVBNWGRGPIGAAHREYVRVG
DGLYLVTSEVPLAGHSGETLDSRGLPFIACHSSCFRRASRSGKDTGTIRVNOF
DKVEGFYCTPADEAHEHERLGMQOMLARIIVPYRVDAAGDGSSAARFDECA
WIPQAGYRELTSNCTTFQARLATRYRDASGKQIAATLTGTLATRMVAITLEN
HORPDGSRVPADLVFVGVELEPVA"
complement(3067..3141)
/gene="sers"
/note="P500179 Aminoacyl-transfer RNA synthetases
class-II signature 1"
4039..5388
/gene="Rv3835"
4039..5388
/gene="Rv3835"
/note="Rv3835, (MTCY01A6.34c), len: 449. Unknown."
/codon_start=1
/translate=1
/product="hypothetical protein Rv3835"
/protein_id="CAB06206.1"
/db_xref="GI:1781122"
/db_xref="SPTREMBL:P96243"
/translation="MLDAPEQDPVDPGDPASPPHGEADQPLPGPRMPRLRASATRRA
LLTLALGGLIAGLVTAIPAVGRAPERLAGYIASNPVPSGAKINASFNVASGDCIM
WPDGTPEASIAIVSCADEHREVAESIDMRTFPGMEYQONAPSPARIQOISEEOCA
AVRRYVGTKEFDPNSKFTISMLWPGDAMRQAGRMVCGLSGPPNQKAFPGKVD
IDOSKVPACTGCGIDATTNQPIDVDVDCAPAMASGVGNVAERPDALPSEPRD
GFTKDACTRMTDYLAPLKRITTLTILITLTPSWSASSRVVASCIGTLGNGGA
TLVNASGALLINGOPVPEPPDIPEERLNIPLIDLPTRPAPPAQOLSTSPSTGH
LPAQPVPTPTRPESHAPASAPASTOPPPAGAPPAQOSEATPPGPAEPAPAG"
5393..5806
/gene="Rv3836"
5393..5806
/gene="Rv3836"
/note="Rv3836, (MTCY01A6.33c), len: 137. Unknown;
containsP500142 Neutral zinc metalloproteinases,
zinc-binding region signature"
/codon_start=1
/translate=1
/product="hypothetical protein Rv3836"
/protein_id="CAB06205.1"
/db_xref="GI:1781121"
/db_xref="SPTREMBL:P96242"
/translation="MTVYRMDPQRFDELVSADLILPELADAMDNVVLVANRHPQHE
NIGQYRGVALTERGSDYAGSLPDATTIREALLDCDSDEYVDQVAITVIEVANH
FGIDDERLDQGRWDEPAPRGKNPDISAPDANNGP"
5672..5701
/gene="Rv3836"
/note="P500142 Neutral zinc metalloproteinases,
zinc-binding region signature"
complement(6001..6699)
/gene="Rv3837c"
complement(6001..6699)

```

Alignment Scores:			
Pred. No.:	2,08e-05	Length:	37751
Score:	94.00	Matches:	15
Percent Similarity:	89.29%	Conservative:	10
Best Local Similarity:	53.57%	Mismatches:	3
Query Match:	64.83%	Indels:	0
DB:	1	Gaps:	0

```

US-09-987-190-2 (1-30) x MTCY1A6 (1-37751)
Oy 1 LysTYSerLeuPProGluLeuAspTY-GluuphSeGrlarhrgluProTYrlsSercly 20
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

Db 15702 GAATACACCTTGCCAGACCTGAGACTACGAGACCTGGAACCGACATCTCGGCT 15761

Qy 21 Glnleangluile\*\*TyrThr 28  
|||||:::|:::|

Db 15762 CAGATCAAGAGCTTACCACACAGC 15785

RESULT 21

PFSSOD 567 bp DNA linear BCT 14-MAR-1996

LOCUS P.freudenreichii sod gene (partial).

DEFINITION X01650

ACCESSION X01650.1 GI:1061201

VERSION X01650

KEYWORDS sod gene; superoxide dismutase.

SOURCE Propionibacterium freudenreichii.

ORGANISM Propionibacterium freudenreichii  
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
Actinomycetales; Propionibacteriaceae; Propionibacteriaceae;  
Propionibacterium.

REFERENCE 1 (bases 1 to 567)  
Gabbianelli, R., Battistoni, A., Polizio, F., Carri, M.T., De Martino, A., Meier, B., Desideri, A. and Rotilio, G.  
Metal uptake of recombinant cambialistic superoxide dismutase from Propionibacterium shermanii is affected by growth conditions of host Escherichia coli cells  
Biochem. Biophys. Res. Commun. 216 (3), 841-847 (1995)

TITLE 7488202

JOURNAL MEDLINE

REFERENCE 2 (bases 1 to 567)  
Gabbianelli, R.  
Direct Submission  
Submitted (19-SEP-1995) R. Gabbianelli, Dept. Biology, University for Vergata, Via Della Ricerca Scientifica, Rome, ITALY  
revised by submitor 03-OCT-95

REMARK FEATURES

source location/Qualifiers

1. 567  
/organism="Propionibacterium freudenreichii"  
/strain="P23"  
/sub\_species="shermanii"  
/db\_xref="taxon:1744"  
1. 567  
/gene="sod"  
1. .>567  
/EC\_number="1.15.1.1"  
/codon\_start=1  
/transl\_table=11  
/product="superoxide dismutase"  
/protein\_id="CAA62838.1"  
/db\_xref="GI:1061202"  
/translation="MAVYTLPELPYDSALEPYSIGEIMELHDKHKAYVDGANTALDKIABARDKADFGAINKLEKDLAFNLAGHNHSPFMKMAPKSGAPRPDEIGAID EFGSGFDNMKAQFTAAATGIGQSGMASLVNDPLGKRINTLQFYDHONNLPAAGSIPLLQ LDMWEHAFYLOYKNVKGDIYKSMNVVNM"

BASE COUNT 118 a 189 c 164 g 96 t

ORIGIN

Alignment Scores:

Pred. No.: 4e-07 Length: 567

Score: 91.00 Matches: 16

Percent Similarity: 84.62% Conservative: 6

Best Local Similarity: 61.54% Mismatches: 4

Query Match: 62.76% Indels: 0

DB: 1 Gaps: 0

US-09-987-190-2 (1-30) x PFSSOD (1-567)

Qy 2 TyserleuprogiluleuapTyrgluPheseralathrgiupProTyriIeserglyGln 21  
|||||:::|:::|

Db 10 TACACGCTGCCGGAACCTCCCTACGACTACTCGCCCTGAGCCGTATACATCTCGGCTAG 69

Qy 22 Ileangluile\*\*Tyr 27  
|||||:::|:::|

Db 70 ATCATGAGACTCCACCAT 87

RESULT 22

PFSSOD 567 bp DNA linear BCT 15-MAY-1998

LOCUS P.freudenreichii sod gene, partial.

DEFINITION Y09012

ACCESSION Y09012.1 GI:3135855

VERSION Y09012

KEYWORDS cytoplasmatic protein; sod gene; superoxide dismutase.

SOURCE Propionibacterium freudenreichii.

ORGANISM Propionibacterium freudenreichii  
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
Actinomycetales; Propionibacteriaceae; Propionibacteriaceae;  
Propionibacterium.

REFERENCE 1 (bases 1 to 567)  
Gabbianelli, R.  
Direct Submission  
Submitted (24-OCT-1996) R. Gabbianelli, Dept. of Biology, University for Vergata, Via Della Ricerca Scientifica, Rome, ITALY  
2 (bases 1 to 567)

REFERENCE 2 (bases 1 to 567)  
Gabbianelli, R., Battistoni, A., Polizio, F., Carri, M.T., De Martino, A., Meier, B., Desideri, A. and Rotilio, G.  
Metal uptake of recombinant cambialistic superoxide dismutase from Propionibacterium shermanii is affected by growth conditions of host Escherichia coli cells  
Biochem. Biophys. Res. Commun. 216 (3), 841-847 (1995)

TITLE 7488202

JOURNAL MEDLINE

REFERENCE 2 (bases 1 to 567)  
Gabbianelli, R.  
Direct Submission  
Submitted (24-OCT-1996) R. Gabbianelli, Dept. of Biology, University for Vergata, Via Della Ricerca Scientifica, Rome, ITALY  
2 (bases 1 to 567)

REMARK FEATURES

source location/Qualifiers

1. 567  
/organism="Propionibacterium freudenreichii"  
/strain="P23"  
/sub\_species="shermanii"  
/db\_xref="taxon:1744"  
1. 567  
/gene="sod"  
1. .>567  
/EC\_number="1.15.1.1"  
/codon\_start=1  
/transl\_table=11  
/product="superoxide dismutase"  
/protein\_id="CAA70215.1"  
/db\_xref="GI:3135856"  
/translation="MAVYTLPELPYDSALEPYSIGEIMELHDKHKAYVDGANTALDKIABARDKADFGAINKLEKDLAFNLAGHNHSPFMKMAPKSGAPRPDEIGAID EFGSGFDNMKAQFTAAATGIGQSGMASLVNDPLGKRINTLQFYDHONNLPAAGSIPLLQ LDMWEHAFYLOYKNVKGDIYKSMNVVNM"

BASE COUNT 118 a 189 c 164 g 96 t

ORIGIN

Alignment Scores:

Pred. No.: 4e-07 Length: 567

Score: 91.00 Matches: 16

Percent Similarity: 84.62% Conservative: 6

Best Local Similarity: 61.54% Mismatches: 4

Query Match: 62.76% Indels: 0

DB: 1 Gaps: 0

US-09-987-190-2 (1-30) x PFSSOD (1-567)

Qy 2 TyserleuprogiluleuapTyrgluPheseralathrgiupProTyriIeserglyGln 21  
|||||:::|:::|

Db 10 TACACGCTGCCGGAACCTCCCTACGACTACTCGCCCTGAGCCGTATACATCTCGGCTAG 69

Qy 22 Ileangluile\*\*Tyr 27  
|||||:::|:::|

Db 70 ATCATGAGACTCCACCAT 87

RESULT 23

LOCUS AB079877 655 bp mRNA linear INV 19-FEB-2002

DEFINITION	Marsupenaeus japonicus sod-1 mRNA for superoxide dismutase like protein, complete cds.
ACCESSION	AB079877
VERSION	AB079877.1 GI:18700488
KEYWORDS	
SOURCE	
ORGANISM	Marsupenaeus japonicus CDNA to mRNA. Marsupenaeus japonicus Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca; Eumalacostraca; Eucarida; Decapoda; Dendrobranchiata; Penaeoidea; Penaeidae; Marsupenaeus.
REFERENCE	1 Soma,G., Inagawa,H., Nishizawa,T., Honda,T., Nomura,Y., Uenohe,M., Takahashi,Y., Itami,T. and Yokomizo,Y. Kuruma shrimp CDNA similar to superoxide dismutase protein Published Only in Database (2002) 2 (bases 1 to 655) Soma,G., Inagawa,H., Nishizawa,T., Honda,T., Nomura,Y., Uenohe,M., Takahashi,Y., Itami,T. and Yokomizo,Y. Direct Submission Submitted (15-FEB-2002) Hitoyuki Inagawa, Tokushima Bunri University, Institute for Health Sciences, Yamashiro-cho Nishihama-bouji, Tokushima, Tokushima 770-8514, Japan (E-mail:pinae@okushima.bunri-u.ac.jp, Tel:81-88-622-9611(ex.1104), Fax:81-88-622-3217)
COMMENT	This study was supported by Grant-in-Aid of Recombinant Cytokine's Project provided by the Ministry of Agriculture, Forestry and Fisheries, Japan (RCP2001-2230). Location/Qualifiers 1..655 /organism="Marsupenaeus japonicus" /db_xref="taxon:27405" 1..655 /gene="sod-1" 32..637 /gene="sod-1" /codon_start=1 /product="superoxide dismutase like protein" /protein_id="BAB85211.1" /db_xref="GI:18700489" /translation="MAEKAEVYISLEKKAELTGIEVDQIKKQFANADEAVAIREE MATVEGIVVOAGVAGTVPQIAQFKFAIINELIGERRAHLPPLKYDFNLPEPRL ISGMIETHTKHGGYINLLIATKKLVEAANDVANMALPAIKFNCGHLNHT IFMTMAPDAGEPGQAIAQALIDSLDSVFQGQPSASAVE"
FEATURES	
source	173 a 156 c 169 g 157 t
gene	
CDS	
BASE COUNT	173 a 156 c 169 g 157 t
ORIGIN	
Alignment Scores:	
Pred. No.:	4,76e-05 Length: 655
Score:	80.00 Matches: 15
%Ident Similarity:	77.78% Conservative: 6
Best Local Similarity:	55.56% Mismatches: 6
Query Match:	55.17% Indels: 0
DB:	Gaps: 3
US-09-987-190-2 (1-30) x AB079877 (1-655)	
Oy	2 TyrSerLeuProGluLeuAspTyrGlnPheSerAlaThrGluProTyrIleSerGlyGln 21 :::     :::     :::     :::     :::     :::     :::     :::
Db	290 CATGCTCTGCCCCCTCAAGTACGACTTAATGCCCCGTGAACACACACTTTGGGCATG 149 
Oy	22 Tleanglulle***TyrThr 28 
Db	350 ATCATGAGATTTCATCACACC 370 
RESULT 24	
AB001693	954 bp mRNA linear NAM 14-DEC-1999
LOCUS	Equus caballus Mn-SOD mRNA for manganese superoxide dismutase,
DEFINITION	complete cds.
ACCESSION	AB001693
VERSION	AB001693.1 GI:4589877
KEYWORDS	Mn-SOD; manganese superoxide dismutase.
SOURCE	Equus caballus (strain:Thoroughbred) adult male testis CDNA to

ORGANISM	MRNA.
	Equus caballus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
TITLE	1 (sites)
JOURNAL	Ishida,N., Katayama,Y., Sato,F., Hasegawa,T. and Mukoyama,H.
MEDLINE	The cDNA sequences of equine antioxidant enzyme genes Cu/Zn-SOD
REFERENCE	and Mn-SOD, and these expressions in equine tissues
	J. Vet. Med. Sci. 61 (3), 291-294 (1998)
	9261591
	2 (bases 1 to 954)
	Ishida,N.
	Submitted (07-MAR-1997) Nobuhige Ishida, JRA Equine Research
	Institute, Laboratory of Molecular and Cellular Biology; 321-4,
	Tokamachi-Cho, Utsunomiya, Tochigi 320-0856, Japan
	(E-mail:nishida@center.equinst.go.jp, Tel:81-28-647-0662,
	Fax:81-28-647-0686)
FEATURES	Location/Qualifiers
SOURCE	1..954
	/organism="Equus caballus"
	/strain="Thorougbred"
	/db_xref="taxon:9796"
	/sex="male"
	/tissue_type="testis"
	/dev_stage="adult"
	/note="domestic horse"
	1..954
	/gene="Mn-SOD"
	70..738
	/gene="Mn-SOD"
	/codon_start=1
	/product="manganese superoxide dismutase"
	/protein_id="BAI6922.1"
	/db_xref="GI:4589878"
	/translation="MLCRACSTSRKLVLPALGSLGRORHSLPDLQYDYGALPEPYINA
	QIMQLHSKHAAVNNINLNTKEKYOEALAKGVTAQIALPALPKFGGHINTFTW
	TNISPNNGSGEPKKLDPAIRDRDGSPDKPRFKETAVASAGIQSGKGHLGPNRQRQ
	IVACPNDPLQGTTGLIPLGIDVWEHAVLYQKNVRPDYLKAIMNVTMENVSERYM
	ACKK"
BASE COUNT	235 a 223 c 256 g 240 t
ORIGIN	
Alignment Scores:	
Pred. NO.:	7.59e-05 Length: 954
Score:	80.00 Matches: 13
Percent Similarity:	78.57% Conservative: 9
Best Local Similarity:	46.43% Mismatches: 6
Query Match:	55.17% Indels: 0
DB:	Gaps: 0
US-09-987-190-2 (1-30) x AB001693 (1-954)	
Oy	1 LysTyrSerLeuProGluLeuAspTyrGluHisSerAlaThrGluProTyrIleSerGly 20
	: : : : :           : : : : :           : : : : :
Db	142 AAGCAGACCTCCCGACTTCAGATGATTATGACGCTCTGGAGCCCTACATCAAGCC 201
	: : : : :           : : : : :           : : : : :
Oy	21 GlnIleGlnIle**TyrThr 28
	: : : : :           : : : : :           : : : : :
Db	202 CAGATCATGCAGCTGCACCACAGC 225
RESULT 25	
GSPJ12188	
LOCUS	GSPJ12188 4439 bp DNA linear BCT 21-MAY-2002
DEFINITION	Gordonia sp. kb2 soda gene, ORF1, ORF2 and ORF3 (partial).
ACCESSION	AJ312188
VERSION	AJ312188.1 GI:21104320
KEYWORDS	ATP-binding protein; Mn-superoxide dismutase; ORF1; ORF2; ORF3; soda gene; transmembrane protein.
SOURCE	Gordonia sp. Kb2.
ORGANISM	Gordonia sp. Kb2 Bacteria; Firmicutes; Actinobacteriae; Actinobacteridae; Actinomycetales; Corynebacterinae; Gordoniaceae; Gordonia.



REFERENCE  
AUTHORS  
JOURNAL  
TITLE  
Berekaa, M.M., Department of Institut fuer Mikrobiologie,  
Westfaelische Wilhelms-Universitaet Muenster, Muenster, Germany

REFERENCE  
AUTHORS  
JOURNAL  
TITLE  
Berekaa, M.M., Priefert, H. and Steinbuechel, A.  
Identification and characterization of an extracellular  
Mn-superoxide dismutase encoding gene (sod) of *Gordonia* sp. strain  
KD2: A radical scavenger enzyme produced during rubber degradation  
Unpublished  
3 (bases 1 to 4439)  
Berekaa, M.M.  
Direct Submission  
Submitted (21-MAY-2001) Berekaa M.M., Institut fuer Mikrobiologie,  
Westfaelische Wilhelms-Universitaet, Corrensstrasse 3,  
48149-Muenster, D-48149, GERMANY

FEATURES  
source  
location/Qualifiers  
1..4439  
/organism="Gordonia sp. KD2"  
/strain="KD2"  
/db\_xref="taxon:160824"  
/country="Germany:Muenster"  
/note="rubber degrading bacterium"  
1..640  
/gene="soda"  
1..3  
/gene="soda"  
11..640  
/gene="soda"  
/function="radical scavenger enzyme"  
/codon\_start=1  
/evidence=experimental  
/transl\_table=1  
/product="Mn-superoxide dismutase"  
/protein\_id="CAC85367.1"  
/db\_xref="GI:21104321"  
/translation="MAEYTLADLDYDALEPHISGRIMELHSHKHATYVKANDTL  
EKLAAREDSIAGKVGILSATLSFHLGHTNHSIFWKNSPNGSGEPGLAAITE  
QGGFDKRAHFTAAATLQSGWALIGYPTIGKLVILTQDSGNI PAALIPVWML  
DMMEHAFYLDYONKPDYKAMWVVMADAGEFRSAKQSGSLIIPA"  
complement (725..1519)  
/function="hypothenical ABC transport system"  
/note="ORF1"  
/codon\_start=1  
/transl\_table=1  
/product="hypothenical protein"  
/protein\_id="CAC85368.1"  
/db\_xref="GI:21104322"  
/translation="MSHRSVRLGRCPMTMTVPLRQVPGDSFVHRLMAGTKLYVL  
ILGIMTVLPSWPAIGVAAIVITLALAGIPLGAI PRPMWFGILGIGAFVNSFA  
ILGALIVLRAITLALVASSIIVITTPMADVAPALRLMRPLRLPLVEMWAVA  
IALCMRGLPLIEELAVLRAAHRRLPAKGRSDHPSAMGIMDLITAMSAIARSAE  
MAELTRAGGTGRITAHPARGRSDALALVIVYACTVAVTITL"  
complement (1516..3765)  
/function="hypothenical ABC transport system"  
/note="ORF2"  
/codon\_start=1  
/transl\_table=1  
/product="hypothenical protein"  
/protein\_id="CAC85369.1"  
/db\_xref="GI:21104323"  
/translation="MCWYPSAGSIRTAEPSCDARLHAIRPTGTRGRSPAGPAADT  
RHGPVPLASAGMEFASVAFGLTVAAVIVASVIRHAGRAARSPRDPGPA  
ARALITATVTAFAFAMAGTAMATVLSALIGYVDIKRRRGSGALGATVAV  
PVLGVSIVLIVLVPRLTITAMTNSLHGKAKMLHKEPPTDVGKALMSRSTIS  
HMMIWMVSGIVGTALSLVVAWMTLGAASSPACPARAHLDGGRDVAAPLARS  
NTSLLOPGARHVLHDLIDPGRVAVANGSGKSTLKLAGRPSTGVHRLP  
AGLVHGTALVQRPETOMIGSRVADVWGLPGADVVEALIAVGLSGDRET  
TDLSGQOORALIAALARPRLIADETSMVDEGRGOLLALPVRGIAVYL  
ITHRGSEAAADRVYIHLAAGRVPHPRMMDPDTDIRPAPADSSATSLAAGPHV  
DOLAARPHLGGPLVLDHYGVYALPSSPWEVALTIDICLTVRGEBLIVGAGSK  
TTLAMVAGLSLPSGTCHLLESSGFTPAAGPDPGVTDHVGAVGLGQHALQIQ  
KTVVDEIMAAAGEKVGTAETAVRLVGLPRQAAATKVDLLSGQMRVYVLAELIAR

HPYMLVINEPLAGINPARRRSVRLPALRADQMTIIVISHDFESLDVCTRRVLY  
DGAVAAQDP"  
complement (1524..1526)  
/note="ORF1"  
complement (3772..3773)  
/note="ORF2"  
3819..3822  
/note="ORF3"  
3832..38439  
/function="unknown"  
/note="ORF3"  
/codon\_start=1  
/transl\_table=1  
/product="hypothenical protein"  
/protein\_id="CAC85370.1"  
/db\_xref="GI:21104324"  
/translation="MOSIPLETPFOHEGLNGLISGRPESTLEWTHSVGVRG  
CRPGLPTITLIVGVRPARSADTVGMMVAISNTTDSFVAGRLNPPALVHL  
PRSRFTLCCGVFGRLVLAAGPTELEHRAWGEATHRELARVGIIDPVPTKV  
LACSSISGPAPKCPGSTRSARQLGPDSDVSGRCSRSLILV"

BASE COUNT 771 a 1508 c 1524 g 636 t

ORIGIN

Alignment Scores:  
Pred. No.: 0.000511 Length: 4439  
Score: 80.00 Matches: 13  
Percent Similarity: 85.71% Conservative: 11  
Best local Similarity: 46.43% Mismatches: 4  
Query Match: 55.17% Indels: 0  
DB: 1 Gaps: 0

US-09-987-190-2 (1-30) x GSP312188 (1-4439)

Qy 1 LysTYrSerLeuProGluLeuAspTYrGluPheSerAlaThrGluProTYrIleSerGly 20  
Db 17 GAATACAGCGCTTCCGATCTGGATTACGACTACCGGACTCGGACCGACATCTCCGCG 76

Qy 21 GlnIleAsnGluIle\*\*TYrThr 28  
Db 77 AGATCATGAGCTCCACACACG 100

RESULT 26  
E15569  
LOCUS 812 bp, DNA linear PAT 28-JUN-1999  
DEFINITION cDNA encoding Malassezia antigen protein MF-4.  
ACCESSION E15569  
VERSION E15569.1 GI:5710252  
KEYWORDS UP 1998077296-A/4.  
SOURCE unidentified.  
ORGANISM unidentified.  
REFERENCE 1 (bases 1 to 812)  
AUTHORS Takesako, K., Daimon, H., Kuroda, M., Katou, I., Yasueda, H., Akiyama, K.  
and Yamaguchi, H.  
RECOMBINANT MALASSEZIA ANTIGENIC PROTEIN AND ITS GENE  
JOURNAL Patent: JP 1998077296-A 4 24-MAR-1998;  
TAKARA SHUZO CO LTD  
OS Malassezia furfur  
PN UP 1998077296-A/4  
PF 05-SEP-1998 JP 1996257613  
PI TAKESAKO KAZUTADA, DAIMON HISASHI, KURODA MASANOBU, PI KATOU  
IKUNOSHIN,  
PI YASUEDA HIROSHI, AKIYAMA KAZUO, YAMAGUCHI HIDEYO PC  
C07K14/195, A61K39/35, C07H21/04, C07K7/08, C12N15/09, C12Q1/68, PC  
G01N33/53,  
PC G01N33/569;  
CC strandedness: Double;  
CC topology: Linear;  
CC hypothetical: No;  
CC anti-sense: No;  
FH Key location/Qualifiers

FT source 1. .812  
FT /strain='TIMM2782', /organism='Malassezia furfur' FT  
FT mac\_peptide 2. .673 /clone='pMFA-4'  
FT /product='Malassezia antigen protein MF-4',  
FEATURES  
source  
1. .812  
/organism='unidentified'  
/db\_xref='taxon:32644'  
BASE COUNT 235 a \_245 c 218 g 114 t  
ORIGIN

Alignment Scores:  
Pred. No.: 9.45e-05 Length: 812  
Score: 79.00 Matches: 15  
Percent Similarity: 70.37% Conservative: 4  
Best Local Similarity: 55.56% Mismatches: 8  
Query Match: 54.48% Indels: 0  
DB: 6 Gaps: 0

US-09-987-190-2 (1-30) x E15569 (1-812)

Qy 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20  
Db 74 AAGTACACGCTCCGCCGCTGCTACGATCAGCGCGCTCGAGCCGCGATCTCGGCG 133

Qy 21 GlnIleAsnGluIle\*\*\*Tyr 27  
Db 134 GAGATCATGAGACGACATC 154

RESULT 27  
AX123311 600 bp DNA linear PAT 11-MAY-2001  
LOCUS  
DEFINITION Sequence 3227 from Patent EP1108790.  
ACCESSION AX123311  
VERSION AX123311.1 GI:14040799  
KEYWORDS  
SOURCE Corynebacterium glutamicum.  
ORGANISM Corynebacterium glutamicum.  
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
Actinomycetales; Corynebacteriaceae; Corynebacteriaceae;  
Corynebacterium.  
1 (bases 1 to 600)  
Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,  
Yokoi, H., Tareishi, N., Senoh, A., Ikeda, M. and Ozaki, A.,  
TITLE Novel polynucleotides  
JOURNAL Patent: EP 1108790-A 3227 20-JUN-2001;  
KYOWA HAKKO KOGYO CO., LTD. (JP)  
FEATURES  
source  
1. .600  
/organism='Corynebacterium glutamicum'  
/db\_xref='taxon:1718'  
BASE COUNT 143 a 197 c 146 g 114 t  
ORIGIN

Alignment Scores:  
Pred. No.: 0.00015 Length: 600  
Score: 77.00 Matches: 13  
Percent Similarity: 74.07% Conservative: 7  
Best Local Similarity: 48.15% Mismatches: 7  
Query Match: 53.10% Indels: 0  
DB: 6 Gaps: 0

US-09-987-190-2 (1-30) x AX123311 (1-600)

Qy 2 TyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 21  
Db 10 TACGAATCCCGAATCGATACGATACGATACGATACGATACGATACGATACGATACG 69

Qy 22 IleAsnGluIle\*\*\*TyrThr 28  
Db 70 ATCATGAGACTTCACCATCTC 90

RESULT 28  
RABMSD 606 bp mRNA linear MAM 11-FEB-1994  
LOCUS  
DEFINITION Oryctolagus cuniculus manganese superoxide dismutase mRNA, partial  
cds.  
ACCESSION L28808  
VERSION L28808.1 GI:454174  
KEYWORDS manganese superoxide dismutase.  
SOURCE Oryctolagus cuniculus (strain NZ white) cDNA to mRNA.  
ORGANISM Oryctolagus cuniculus  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
REFERENCE  
AUTHORS Jackson, R.M.  
JOURNAL Unpublished (1994)  
FEATURES  
source  
Location/Qualifiers  
1. .606  
/organism='Oryctolagus cuniculus'  
/strain='NZ white'  
/db\_xref='taxon:9986'  
<1. .>606  
/codon\_start=1  
/product='manganese superoxide dismutase'  
/protein\_id='AA31401.1'  
/db\_xref='GI:454175'  
/translation='HGRGMKSLPDLPYDYGALEPHINAIQIMELHSHKHHAAVYNNLN  
ATEBKREALARGDVTAAHVALOPALFKGGKHNTIPMTNISPNGSGEPKGLLEAI  
KRDPSGPRKRELTAIVSVVGSGMGWGLFKNKECHLOIAACANDPLOGTGLPL  
LGIDVMAHYLQIKVRPDYLAIKMNVITMENVTERIMACK'  
BASE COUNT 155 a 162 c 178 g 111 t  
ORIGIN

Alignment Scores:  
Pred. No.: 0.000152 Length: 606  
Score: 77.00 Matches: 13  
Percent Similarity: 78.57% Conservative: 9  
Best Local Similarity: 46.43% Mismatches: 6  
Query Match: 53.10% Indels: 0  
DB: 4 Gaps: 0

US-09-987-190-2 (1-30) x RABMSD (1-606)

Qy 1 LysTyrSerLeuProGluLeuAspTyrGluPheSerAlaThrGluProTyrIleSerGly 20  
Db 16 AAGCACAGCTCCCGACCTGCTCCTACGATCAGCGCGCTGAGCCACATCAACGCG 75

Qy 21 GlnIleAsnGluIle\*\*\*TyrThr 28  
Db 76 CAGATCATGAGACTGCACACGCG 99

RESULT 29  
AB055218 960 bp DNA linear BCT 02-AUG-2001  
LOCUS  
DEFINITION Corynebacterium glutamicum gene for superoxide dismutase, complete  
cds.  
ACCESSION AB055218  
VERSION AB055218.1 GI:15076595  
KEYWORDS  
SOURCE Corynebacterium glutamicum (strain:ATCC 13032) DNA.  
ORGANISM Corynebacterium glutamicum  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.  
1  
Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Hattori, M.,  
Shiba, T., Sakaki, Y., Yokoi, H. and Ozaki, A.,  
TITLE SOD of Corynebacterium glutamicum ATCC 13032  
JOURNAL Published Only in Databases (2001)  
REFERENCE  
AUTHORS Nakagawa, S.  
JOURNAL Direct Submission  
Submitted (01-FEB-2001) Satooshi Nakagawa, Kyowa Hakko Kogyo Co.  
Ltd., Tokyo Research Laboratories; 3-6-6, Asahi-machi, Machida,

Tokyo 194-8533, Japan (E-mail: snakagawa@kyowa.co.jp,  
Tel:81-42-725-2555(ex.2165), Fax:81-42-726-8330)

## FEATURES

## source

Location/Qualifiers

## CDS

1. .960  
/organism="Corynebacterium glutamicum"  
/strain="ATCC 13032"  
/db\_xref="taxon:1718"  
154. .756  
/EC\_number="1.15.1.1"  
/codon\_start=1  
/transl\_table=1  
/product="superoxide dismutase"  
/protein\_id="BAB62412.1"  
/db\_xref="GI:15076596"  
/translation="MAVVELPELDYADALEPHIAIMELHSHKHATYVAGANAL  
EALAKAREGTPQIRALSKNLAFNIGTHSVFPMKLSPNGGEPTELAIR  
DPSFAFQDHPNSAALGLOGSNAVIGYDHISGRLLVIBOLTDOQNIISVDITFVLM  
DMEMHAYLYQKVKADYKAVVNMVFWDDAAAFPAASK"

## BASE COUNT

231 a 278 c 236 g 215 t

## Alignment Scores:

Pred. No.: 0.000268 Length: 960  
Score: 77.00 Matches: 13  
Percent Similarity: 74.07% Conserved: 7  
Best Local Similarity: 48.15% Mismatches: 7  
Query Match: 53.10% Indels: 0  
DB: 1 Gaps: 0

US-09-987-190-2 (1-30) x AB055218 (1-960)

Qy 2 TyserleuprogulueapTyrgluPheserAathrgluProTyrlIseerGlyGln 21

Db 163 TACGAAGTCCAGAACTGACATACGATACGACGCTCTCGAGCCACACATCGCGCTGAA 222

Qy 22 Ileangluile\*\*TyrrThr 28

Db 223 ATCATGAGCTTCACACTCC 243

## RESULT 30

AX136077 1143 bp DNA linear PAT 30-MAY-2001

LOCUS AX136077

DEFINITION Sequence 1 from Patent EP1077261.

ACCESSION AX136077

VERSION AX136077.1 GI:14272487

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## DEPOSIT

## FEATURES

## source

Location/Qualifiers

1. .1143

/organism="Corynebacterium melassecola"

/db\_xref="taxon:41643"

/note="ATCC 17965"

338. 940

/note="unnamed protein product"

/codon\_start=1

/transl\_table=1

/protein\_id="CAC39682.1"

/db\_xref="GI:14272488"

/translation="MAVVELPELDYADALEPHIAIMELHSHKHATYVAGANAL  
EALAKAREGTPQIRALSKNLAFNIGTHSVFPMKLSPNGGEPTELAIR  
DPSFAFQDHPNSAALGLOGSNAVIGYDHISGRLLVIBOLTDOQNIISVDITFVLM  
DMEMHAYLYQKVKADYKAVVNMVFWDDAAAFPAASK"

BASE COUNT 273 a 323 c 288 g 259 t  
ORIGIN

## Alignment Scores:

Pred. No.: 0.000333 Length: 1143  
Score: 77.00 Matches: 13  
Percent Similarity: 74.07% Conserved: 7  
Best Local Similarity: 48.15% Mismatches: 7  
Query Match: 53.10% Indels: 0  
DB: 6 Gaps: 0

US-09-987-190-2 (1-30) x AX136077 (1-1143)

Qy 2 TyserleuprogulueapTyrgluPheserAathrgluProTyrlIseerGlyGln 21

Db 347 TACGAAGTCCAGAACTGACATACGACGCTCTCGAGCCACACATCGCGCTGAA 406

Qy 22 Ileangluile\*\*TyrrThr 28

Db 407 ATCATGAGCTTCACACTCC 427

## RESULT 31

BD011434 1143 bp DNA linear PAT 31-JAN-2002

LOCUS BD011434

DEFINITION DNA, amino acid sequence, coryneform microorganism, shuttle vector,  
method of potentiating superoxide dismutase activity and method of  
producing metabolite.

ACCESSION BD011434

VERSION BD011434.1 GI:18639807

KEYWORDS UP 2001078788-A/1.

SOURCE unidentified.

ORGANISM unidentified.

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

## BASE COUNT

## ORIGIN

## Alignment Scores:

Pred. No.: 0.000333 Length: 1143  
Score: 77.00 Matches: 13  
Percent Similarity: 74.07% Conserved: 7  
Best Local Similarity: 48.15% Mismatches: 7  
Query Match: 53.10% Indels: 0  
DB: 6 Gaps: 0

US-09-987-190-2 (1-30) x BD011434 (1-1143)

Qy 2 TyserleuprogulueapTyrgluPheserAathrgluProTyrlIseerGlyGln 21

Db 347 TACGAAGTCCAGAACTGACATACGACGCTCTCGAGCCACACATCGCGCTGAA 406





```

/codon_start=1
/transl_table=1
/product="Hypothetical protein"
/protein_id="BAC00200.1"
/db_xref="GI:21325579"
/translation="MTVSHALGFKRRQESLELSLRSNFPVLAVVAQYFQGAIA
AKPASELYQLSDRFLVREGEFELPKSPSYVDWVKSVMFVRPSSQGEVRES
EELAVLDSVORMNPRIASRIESTLQALALESDPTAKRLALEERDRLE
ROIYAVHAGEFEVLTVOIGDRAVIDLIALSIDPARVHELSDNRLRLQDLDP
EDSRDVLIEEFVVDLIGSDACRSPSPFDVLIDERSLIDRMITREVLGRDL
LDSKRTGLRIFRDMEDASFVNGEMTGLARLRHYTTTEFPAESRMIDLDITS
AAALAAEAGEVTSLNHMDTPIVRIGMDVRSIAGLKNPGEERVEDLPEVECELDT
EVLMEQIRASEIDPELEAVSLVLAQSHATITREVLHPFATQGLASIVGLVLAAR
DGVPTRAOIVMESDATHRRITGMQFIRGLNSDLAESEMDK"
/complement(11995..13032)
/gene="Cg12807"
/complement(11995..13032)
/gene="Cg12807"
/EC_number="1.1.1.1"
/nc="PF00107:Zinc-binding dehydrogenases
TIGR00692:cdh: L-threonine 3-dehydrogenase"
/codon_start=1
/transl_table=1
/product="Zn-dependent alcohol dehydrogenases"
/protein_id="BAC00201.1"
/db_xref="GI:21325580"
/translation="MTTAAPOEFTAAVVERKGDVTYKIDLPKRGPHQALVKTLSG
ICHTDLALBEDMPVKRPPRPVPHGEGVELEGREHDVYKQDIDYGNMISACCT
CEYCTIGRETCNAEYGGYTONGSFGGYMLVTRVARIPIGVGYLEBAATLCAGT
VYKALKVSERPGQFVWISGVGLGHIAVQAAAMGRVIAVDIADKLELARGHAE
FTVARNEDSGEAVOKYRTNGAGVGLVAVAEAAFGQALDMARAGTIVFNGLPGEF
PASYNIYFKGLTRGSLVGTRODLAALDFEAGLTKPTVSECSLDEVNGVLDRMEN
GKIDGRAVIRE"
13402..13520
/gene="Cg12808"
13402..13520
/gene="Cg12808"
/codon_start=1
/transl_table=1

```

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:
0.368	77.00	325651	13
Percent Similarity:	74.07%	Conservative:	7
Best Local Similarity:	48.15%	Mismatches:	7
Query Match:	53.10%	Indels:	0
DB:	1	Gaps:	0

-09-987-190-2 (1-30) x AP005283 (1-325651)

2 TyrsrlauprogulueaspyrtyglupheserlathrgluProtyrTlleSerglygin 21  
 Db 142652 TACGAACCTCCGAACTCCGACATACGACGCTCTCGACACACATCCGCCCTGAA 142711

Oy 22 lleaenglulle\*\*TyrThr 28  
 Db 142712 ATCATGAGCTTACACACTCC 142732

RESULT 35  
 AF521909 663 bp mRNA linear INV 07-JUL-2002  
 LOCUS Trichinella pseudospiralis Mn superoxide dismutase mRNA, complete  
 DEFINITION cde.  
 ACCESSION AF521909  
 VERSION AF521909.1 GI:21702728  
 SOURCE Trichinella pseudospiralis.  
 ORGANISM Eukaryota; Metazoa; Nematoda; Enoplea; Trichocephalida;  
 Trichinellidae; Trichinella.  
 REFERENCE 1 (bases 1 to 663)  
 AUTHORS Wu,H.W.K., Ko,R. and Mak,C.H.  
 TITLE Characterization and molecular cloning of Mn- superoxide dismutase

JOURNAL in Trichinella pseudospiralis  
 REFERENCE 2 (bases 1 to 663)  
 AUTHORS Wu,H.W.K., Ko,R. and Mak,C.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (16-JUN-2002) Zoology, The University of Hong Kong, Pokfulam, Hong Kong, China  
 FEATURES  
 source  
 1..663  
 /organism="Trichinella pseudospiralis"  
 /db\_xref="taxon:6337"  
 CDS  
 1..663  
 /db\_start=1  
 /product="Mn superoxide dismutase"  
 /protein\_id="AA76074.1"  
 /db\_xref="GI:21702729"  
 /translation="MALFRLKSPAIRMOCKRKHLPDLPVDFGALPEYIAEIMPL  
 HHQKHHTYNNLNINAEKREKELAKDITTAIALPALPFRNGGHHNHSIFKNILAP  
 KSGGMPFGKLAQAITNNFGSFDI FKSALTAIAVGGSGWAMGIDANMRRLQIATCA  
 NODPLQATTPGIPILGIDWEHAYVYQYKVRGDIYKAIWEVINMDVQGVADAVGS  
 GN"

BASE COUNT 178 a 147 c 162 g 176 t  
 ORIGIN

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:
0.000258	76.00	663	13
Percent Similarity:	74.07%	Conservative:	7
Best Local Similarity:	48.15%	Mismatches:	7
Query Match:	52.41%	Indels:	0
DB:	3	Gaps:	0

US-09-987-190-2 (1-30) x AF521909 (1-663)

Oy 1 LysTyrsrlauprogulueaspyrtyglupheserlathrgluProtyrTlleSergly 20  
 Db 58 AAGACACGTTACCCGATTTGCGGTATGACTTCGATCGTGAACCGTACATCTGCT 117

Oy 21 Glulleaenglulle\*\*Tyr 27  
 Db 118 GAATTTATGCGATTGATCAT 138

RESULT 36  
 DLA278864 761 bp mRNA linear PLN 24-AUG-2000  
 LOCUS Digitalis lanata partial mRNA for manganese superoxide dismutase  
 DEFINITION (Mn-SOD gene).  
 ACCESSION DLA278864  
 VERSION A0278864.1 GI:9929160  
 SOURCE Digitalis lanata.  
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Asteridae; euasterids I; Lamiales; Veroniceae; Digitalis.  
 REFERENCE 1 (bases 1 to 761)  
 AUTHORS Gruner,B., Reva,V.A. and Mueller-Urli,F.  
 TITLE Stress-induction of Mn-dependent SOD and CAT in Digitalis lanata  
 (Ehrh.) during somatic embryogenesis  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 761)  
 AUTHORS Gruner,B.  
 TITLE Direct Submission  
 JOURNAL Submitted (23-AUG-2000) Gruner B., Institute of Pharmaceutical  
 Biology, University of Halle, Hoher Weg 8; Halle /Saale, D-06120,  
 GERMANY

FEATURES  
 source  
 1..761  
 /organism="Digitalis lanata"  
 /db\_xref="taxon:49450"  
 /tissue\_1ib="leaves"  
 1..652  
 /gene="MN-SOD"

## CDS

<1. .652  
 /gene="MN-SOD"  
 /BC number="1.15.1.1"  
 /codon start=2  
 /product="manganese superoxide dismutase"  
 /protein\_id="CAC05260.1"  
 /db\_xref="GI:9929161"  
 /translation="FELPAQAMASRELQKISLPDLPDYAALEPVIAGEIMQITHQKH  
 HOAYITNNALAQIDLSAIGYDAATVVKLHNKIKFNGGHVNSIFWENLAVSOGG  
 GEPPLGPIABIDIKHFGSLBALIKNNAEAGVGSGMWLGVDKOLKRLVYETTANQ  
 DVLATKGASLIVPLGIDVWEHNAVYLOKARPRYLKINIKMMWKTASRYQKECP"

BASE COUNT 219 a 172 c 181 g 189 t

## Alignment Scores:

Pred. No.: 0.000306 Length: 761  
 Score: 76.00 Matches: 14  
 Percent Similarity: 77.78% Conservative: 7  
 Local Similarity: 51.85% Mismatches: 6  
 Indels: 0  
 Gaps: 0

US-09-987-190-2 (1-30) x DLA278864 (1-761)

Qy 1 LysTySerLeuProGluLeuAspTYrGluPheSerAlaThrGluProTYrIleSerGly 20  
 Db 44 AAGATTTCGCTGCCCGCATCTGCCTTACGATTACGCCGCTTGAGCCAGTATCAGCGGC 103  
 Qy 21 GlnIleAsnGluIle\*\*\*TYr 27  
 Db 104 GAGATCATGCAATTCACAC 124

RESULT 37  
 AX469763  
 LOCUS AX469763 325 bp DNA linear PAT 16-JUL-2002  
 DEFINITION Sequence 11 from Patent WO0240498.  
 ACCESSION AX469763  
 VERSION AX469763.1 GI:21901883  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1  
 Obery, L.W., Weydert, C.J. and Smith, B.B.  
 Reduction of antioxidant enzyme levels in tumor cells using  
 antisense oligonucleotides  
 Patent: WO 0240498-A 11 23-MAY-2002;  
 UNIVERSITY OF IOWA RESEARCH FOUNDATION (US)

FEATURES  
 source 1..325  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"

BASE COUNT 76 a 99 c 98 g 52 t

## Alignment Scores:

Pred. No.: 0.000246 Length: 325  
 Score: 74.00 Matches: 12  
 Percent Similarity: 78.57% Conservative: 10  
 Local Similarity: 42.86% Mismatches: 6  
 Indels: 0  
 Gaps: 0

US-09-987-190-2 (1-30) x AX469763 (1-325)

Qy 1 LysTySerLeuProGluLeuAspTYrGluPheSerAlaThrGluProTYrIleSerGly 20  
 Db 113 AAGCAGACGCTCCCGCATCTGCCTTACGATTACGCCGCTTGAGCCAGTATCAGCGGC 172  
 Qy 21 GlnIleAsnGluIle\*\*\*TYrThr 28  
 Db 173 CAGATCATGCAATTCACAC 196

RESULT 38

S78832S1

535 bp

DNA

linear

ROD 24-OCT-1995

LOCUS S78832S1

Sod-2=manganese superoxide dismutase [mice, C3H/HeJ, Genomic, 535  
 nt, segment 1 of 4].

DEFINITION

578832  
 S78832.1 GI:10372116

ACCESSION

1 of 4  
 Mus sp. C3H/HeJ.

VERSION

KEYWORDS

SEGMENT

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 535)  
 Disilver, D., Kleeberger, S.R., Johns, J. and Levitt, R.C.  
 Structure and DNA sequence of the mouse MnSOD gene  
 Mamm. Genome 6 (4), 281-284 (1995)

JOURNAL

MEDLINE

PUBMED

REMARK

Genbank staff at the National Library of Medicine created this  
 entry [NCBI gblseq 169396] from the original journal article.  
 This sequence comes from Fig. 1.  
 Map location: 17.

FEATURES

source 1..535  
 Location/Qualifiers  
 /organism="Mus sp."  
 /db\_xref="taxon:10095"

BASE COUNT 96 a 166 c 171 g 99 t 3 others

ORIGIN

Alignment Scores:  
 Pred. No.: 0.000456 Length: 535  
 Score: 74.00 Matches: 12  
 Percent Similarity: 78.57% Conservative: 10  
 Local Similarity: 42.86% Mismatches: 6  
 Indels: 0  
 Gaps: 0

US-09-987-190-2 (1-30) x S78832S1 (1-535)

Qy 1 LysTySerLeuProGluLeuAspTYrGluPheSerAlaThrGluProTYrIleSerGly 20  
 Db 210 AAGCAGACGCTCCCGCATCTGCCTTACGATTACGCCGCTTGAGCCAGTATCAGCGGC 269

Qy 21 GlnIleAsnGluIle\*\*\*TYrThr 28  
 Db 270 CAGATCATGCAATTCACAC 293

RESULT 39

AR130379

594 bp

DNA

linear

PAT 16-MAY-2001

LOCUS AR130379

DEFINITION Sequence 1 from patent US 6190658.  
 ACCESSION AR130379  
 VERSION AR130379.1 GI:14118704  
 KEYWORDS

ORGANISM

Unknown.

REFERENCE

1 (bases 1 to 594)  
 McCord, J.M., Gao, B. and Flores, S.C.  
 Genetically modified manganese superoxide dismutase for treating  
 oxidative damage

JOURNAL

PATENT

US 6190658-A 1 20-FEB-2001;

FEATURES

source 1..594  
 Location/Qualifiers  
 /organism="unknown"

BASE COUNT 166 a 144 c 154 g 130 t

ORIGIN

Alignment Scores:  
 Pred. No.: 0.000519 Length: 594  
 Score: 74.00 Matches: 12





GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 9, 2003, 13:56:15 ; Search time 5.09174 Seconds

(without alignments)  
173.357 Million cell updates/sec

Title: US-09-987-190-2

Perfect score: 145

Sequence: 1 KYSLPEIDYFSATEPYISQINEIXYTX 30

Scoring table: BLAST62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PTUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/Backlist.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	141	97.2	30	4	US-09-262-856A-2
2	100	69.0	233	2	US-08-928-692-31
3	100	69.0	233	4	US-09-339-972-31
4	79	54.5	30	4	US-09-091-097-52
5	79	54.5	224	4	US-09-091-097-8
6	74	51.0	198	6	US-09-075-019-2
7	74	51.0	198	6	5240847-10
8	74	51.0	198	6	5240847-11
9	74	51.0	222	2	US-08-365-486A-28
10	74	51.0	222	2	US-08-023-980B-44
11	74	51.0	222	2	US-08-486-953A-52
12	74	51.0	222	2	US-08-927-230A-2
13	74	51.0	222	2	US-08-927-230A-3
14	74	51.0	222	3	US-09-151-052-2
15	74	51.0	222	3	US-09-151-052-3
16	74	51.0	222	4	US-09-126-109-2
17	74	51.0	222	4	US-08-880-342-28
18	74	51.0	226	4	US-09-075-019-7
19	71	49.0	188	4	US-09-262-856A-6
20	70	48.3	46	6	5240847-24
21	69	47.6	230	2	US-08-928-692-30
22	69	47.6	230	4	US-09-339-972-30
23	62	42.8	206	4	US-09-091-097-6
24	58	40.0	214	4	US-09-411-578-1
25	57	39.3	201	4	US-09-134-001C-3584
26	56	38.6	190	4	US-08-679-493A-184
27	53	36.6	204	1	US-08-445-909A-17

28	53	36.6	204	1	US-08-445-909A-29	Sequence 29, Appl
29	50	34.5	371	4	US-09-605-785-708	Sequence 708, App
30	50	34.5	553	4	US-09-020-956-113	Sequence 113, App
31	50	34.5	553	4	US-09-030-607-113	Sequence 113, App
32	50	34.5	553	4	US-09-605-785-113	Sequence 113, App
33	50	34.5	553	4	US-09-429-113-113	Sequence 113, App
34	50	34.5	553	4	US-09-352-616A-113	Sequence 113, App
35	50	34.5	553	4	US-09-602-977A-101	Sequence 113, App
36	50	34.5	553	4	US-09-232-149A-113	Sequence 113, App
37	48	33.1	885	1	US-08-042-747A-8	Sequence 8, Appl
38	48	33.1	885	3	US-08-804-439A-23	Sequence 23, Appl
39	48	33.1	885	3	US-08-720-229-23	Sequence 23, Appl
40	48	33.1	885	3	US-08-446-345-36	Sequence 36, Appl
41	48	33.1	1174	2	US-08-446-345-36	Sequence 4, Appl
42	46	31.7	1009	4	US-09-693-146-4	Sequence 11, Appl
43	45.5	31.4	203	3	US-08-331-625A-11	Sequence 11, Appl
44	45.5	31.4	203	4	US-09-484-151-11	Sequence 11, Appl
45	45.5	31.4	208	4	US-08-896-933-30	Sequence 30, Appl

## ALIGNMENTS

RESULT 1  
US-09-262-856A-2  
Sequence 2, Application US/09262856A  
Patent No. 6333164  
GENERAL INFORMATION:  
APPLICANT: TAKESAKO, Kautech  
APPLICANT: WIZUTANI, Shigetoshi  
APPLICANT: ENDO, Masahiro  
APPLICANT: KATO, Ikunosshin  
TITLE OF INVENTION: FUNGAL ANTIGENS AND PROCESS FOR PRODUCING THE SAME  
FILE REFERENCE: 1422-372P  
CURRENT APPLICATION NUMBER: US/09/262,856A  
CURRENT FILING DATE: 1999-03-04  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 2  
LENGTH: 30  
TYPE: PRT  
ORGANISM: Candida albicans  
FEATURE:  
OTHER INFORMATION: residues 26, 39, 30 = unknown

Query Match 97.2%, Score 141, DB 4, Length 30;  
Best Local Similarity 100.0%, Pred. No. 9.9e-17;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
DB 1 KYSLPEIDYFSATEPYISQINEIXYTX 28  
US-09-262-856A-2  
RESULT 2  
US-08-928-692-31  
Sequence 31, Application US/08928692  
Patent No. 5958727  
GENERAL INFORMATION:  
APPLICANT: Brody, Howard  
APPLICANT: Yaver, Deborah S.  
APPLICANT: Lamma, Michael  
APPLICANT: Hansen, Kim  
TITLE OF INVENTION: Methods for Modifying the Production of  
NUMBER OF SEQUENCES: 80  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 5958727 No. 5958727disk of No. 5958727th America, Inc.  
STREET: 405 Lexington Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA



RESULT 5  
US-09-091-097-8  
; Sequence 8, Application US/09091097  
; Patent No. 6432407  
; GENERAL INFORMATION:  
; APPLICANT: TAKESAKO, KAZUTOH  
; APPLICANT: OKADO, TAKASHI  
; APPLICANT: YAGIHARA, TOMOKO  
; APPLICANT: KURODA, MASANOBU  
; APPLICANT: ONISHI, YOSHIMI  
; APPLICANT: KATO, IKUNOSHIN  
; APPLICANT: AKIYAMA, KAZUO  
; APPLICANT: YASUEDA, HIROSHI  
; APPLICANT: YAMAGUCHI, HIDEYO  
; TITLE OF INVENTION: ANTIGENIC PROTEIN ORIGINATING IN  
; TITLE OF INVENTION: MALASSEZIA  
; NUMBER OF SEQUENCES: 58  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH, LLP  
; STREET: PO BOX 747  
; CITY: FALLS CHURCH  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22040-0747  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/091,097  
; FILING DATE:  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WEINER, MARC S.  
; REGISTRATION NUMBER: 32,181  
; REFERENCE/DOCKET NUMBER: 1422-0346P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-205-8000  
; TELEFAX: 703-205-8050  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 224 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; 09-091-097-8

Query Match 54.5%; Score 79; DB 4; Length 224;  
Best Local Similarity 55.6%; Pred. No. 2.4e-05;  
Matches 15; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 KXSLPDLPEYGFATPEYISQINEIXY 27  
Db 25 KYTLPLPYDYGALPEALSGIMETHY 51

RESULT 6  
US-09-075-019-2  
; Sequence 2, Application US/09075019  
; Patent No. 6190658  
; GENERAL INFORMATION:  
; APPLICANT: UTC IR459  
; TITLE OF INVENTION: A GENETICALLY MODIFIED MANGANESE  
; TITLE OF INVENTION: SUPEROXIDE DISMUTASE FOR TREATING OXIDATIVE DAMAGE  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Sheridan Ross P.C.  
; STREET: 1700 Lincoln St., Suite 3500  
; CITY: Denver  
; STATE: CO

COUNTRY: USA  
; ZIP: 80203  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/075,019  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kovarik, Joseph E.  
; REGISTRATION NUMBER: 33,005  
; REFERENCE/DOCKET NUMBER: 2848-22  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (303) 863-9700  
; TELEFAX: (303) 863-0223  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 198 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; 09-075-019-2

Query Match 51.0%; Score 74; DB 4; Length 198;  
Best Local Similarity 42.9%; Pred. No. 0.00014;  
Matches 12; Conservative 10; Mismatches 6; Indels 0; Gaps 0;

QY 1 KXSLPDLPEYGFATPEYISQINEIXYT 28  
Db 1 KXSLPDLPEYGFATPEYISQINEIXYT 28

RESULT 7  
5240847-10  
; Patent No. 5240847  
; APPLICANT: HECKL, KONRAD; SPEYAK, WALTER; OSTERMANN, ELINBORG;  
; ZOPHEL, ANDREAS; KRYSSTEK, EDELTRAUD; MAURER-FOG, INGRID;  
; WICHE-CASTANON, MARIA J.; STRATOWA, CHRISTIAN; HAUPTMANN, RUDOLF  
; TITLE OF INVENTION: HUMAN MANGANESE SUPEROXIDE DISMUTASE  
; (HMN-SOD)  
; NUMBER OF SEQUENCES: 34  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/167,261  
; FILING DATE: 11-MAR-1988  
; SEQ ID NO: 10:  
; LENGTH: 198  
; 5240847-10

Query Match 51.0%; Score 74; DB 6; Length 198;  
Best Local Similarity 42.9%; Pred. No. 0.00014;  
Matches 12; Conservative 10; Mismatches 6; Indels 0; Gaps 0;

QY 1 KXSLPDLPEYGFATPEYISQINEIXYT 28  
Db 1 KXSLPDLPEYGFATPEYISQINEIXYT 28

RESULT 8  
5240847-11  
; Patent No. 5240847  
; APPLICANT: HECKL, KONRAD; SPEYAK, WALTER; OSTERMANN, ELINBORG;  
; ZOPHEL, ANDREAS; KRYSSTEK, EDELTRAUD; MAURER-FOG, INGRID;  
; WICHE-CASTANON, MARIA J.; STRATOWA, CHRISTIAN; HAUPTMANN, RUDOLF  
; TITLE OF INVENTION: HUMAN MANGANESE SUPEROXIDE DISMUTASE  
; (HMN-SOD)  
; NUMBER OF SEQUENCES: 34  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/167,261  
; FILING DATE: 11-MAR-1988  
; SEQ ID NO: 11:



REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 00786/223002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/428-0200  
TELEFAX: 617/428-7045  
TELEX:  
INFORMATION FOR SEQ ID NO: 52:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 222 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-486-953A-52

Query Match 51.0%; Score 74; DB 2; Length 222;  
Best Local Similarity 42.9%; Pred. No. 0.00016;  
Matches 12; Conservative 10; Mismatches 6; Indels 0; Gaps 0;

1 KXSLPDLDFEFSATEPIYSGQINEIXYT 28  
25 KXSLPDLDFYDYGALPHINQIMQLHHS 52

RESULT 12  
US-08-927-230A-2  
Sequence 2, Application US/08927230A  
Patent No. 5985633  
GENERAL INFORMATION:  
APPLICANT: Nick et al.  
TITLE OF INVENTION: Human Manganese Superoxide Dismutase Mutants  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD, LLP  
STREET: 28 State Street, Floor 24  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1875  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/927,230A  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
ATTORNEY/AGENT INFORMATION:  
FILING DATE:  
NAME: Remillard, Jane  
REGISTRATION NUMBER: 38,872  
REFERENCE/DOCKET NUMBER: UFI-001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)742-4214  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 222 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: OTHER  
LOCATION: 131  
OTHER INFORMATION: Xaa can code for Gln or Glu  
US-08-927-230A-2

Query Match 51.0%; Score 74; DB 2; Length 222;  
Best Local Similarity 42.9%; Pred. No. 0.00016;  
Matches 12; Conservative 10; Mismatches 6; Indels 0; Gaps 0;

QY 1 KXSLPDLDFEFSATEPIYSGQINEIXYT 28  
DB 25 KXSLPDLDFYDYGALPHINQIMQLHHS 52

RESULT 13  
US-08-927-230A-3  
Sequence 3, Application US/08927230A  
Patent No. 5985633  
GENERAL INFORMATION:  
APPLICANT: Nick et al.  
TITLE OF INVENTION: Human Manganese Superoxide Dismutase Mutants  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD, LLP  
STREET: 28 State Street, Floor 24  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1875  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/927,230A  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Remillard, Jane  
REGISTRATION NUMBER: 38,872  
REFERENCE/DOCKET NUMBER: UFI-001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)742-4214  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 222 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: OTHER  
LOCATION: 131  
OTHER INFORMATION: Xaa can code for Gln or Glu  
US-08-927-230A-3

Query Match 51.0%; Score 74; DB 2; Length 222;  
Best Local Similarity 42.9%; Pred. No. 0.00016;  
Matches 12; Conservative 10; Mismatches 6; Indels 0; Gaps 0;

QY 1 KXSLPDLDFEFSATEPIYSGQINEIXYT 28  
DB 25 KXSLPDLDFYDYGALPHINQIMQLHHS 52

RESULT 14  
US-09-151-052-2  
Sequence 2, Application US/09151052  
Patent No. 6107070  
GENERAL INFORMATION:  
APPLICANT: Nick et al.  
TITLE OF INVENTION: Human Manganese Superoxide Dismutase Mutants  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 28 State Street, Floor 24  
CITY: Boston  
STATE: Massachusetts



TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-126-109-2

Query Match 51.0%; Score 74; DB 4; Length 222;  
Best Local Similarity 42.9%; Pred. No. 0.00016;  
Matches 12; Conservative 10; Mismatches 6; Indels 0; Gaps 0;

QY 1 KXSLPELDYERSATEPIYSGQINIEIXYT 28  
DB 25 KXSLPELDYIDYGALEPHINAIQIMQLHHS 52

RESULT 17  
US-08-880-342-28  
Sequence 28, Application US/08880342  
Patent No. 6218179

GENERAL INFORMATION:  
APPLICANT: Webster, Keith A.  
APPLICANT: Bishopric, Nanette H.  
APPLICANT: Murphy, Brian  
APPLICANT: Laderoute, Keith R.  
APPLICANT: Green, Christopher J.  
TITLE OF INVENTION: Tissue Specific Hypoxia Regulated  
TITLE OF INVENTION: Therapeutic Constructs  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dehlinger & Associates  
STREET: 350 Cambridge Avenue, Suite 250  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/880,342  
FILING DATE: 23-JUN-1997  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/IB95/00996  
FILING DATE: 13-NOV-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/365,486  
FILING DATE: 23-DEC-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Sholtz, Charles K.  
REGISTRATION NUMBER: 38,615  
REFERENCE/DOCKET NUMBER: 8255-0018.30  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 222 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-880-342-28

Query Match 51.0%; Score 74; DB 4; Length 222;  
Best Local Similarity 42.9%; Pred. No. 0.00016;  
Matches 12; Conservative 10; Mismatches 6; Indels 0; Gaps 0;

QY 1 KXSLPELDYERSATEPIYSGQINIEIXYT 28  
DB 25 KXSLPELDYIDYGALEPHINAIQIMQLHHS 52

RESULT 18  
US-09-075-019-7  
Sequence 7, Application US/09075019  
Patent No. 6190658

GENERAL INFORMATION:  
APPLICANT: UTC IR459  
TITLE OF INVENTION: A GENETICALLY MODIFIED MANGANESE  
TITLE OF INVENTION: SUPEROXIDE DISMUTASE FOR TREATING OXIDATIVE DAMAGE  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheridan Ross P.C.  
STREET: 1700 Lincoln St., Suite 3500  
CITY: Denver  
STATE: CO  
COUNTRY: USA  
ZIP: 80203

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/075,019  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Kovarik, Joseph E.  
REGISTRATION NUMBER: 33,005  
REFERENCE/DOCKET NUMBER: 2848-22  
TELEPHONE: (303) 863-9700  
TELEFAX: (303) 863-0223  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 226 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-075-019-7

Query Match 51.0%; Score 74; DB 4; Length 226;  
Best Local Similarity 42.9%; Pred. No. 0.00016;  
Matches 12; Conservative 10; Mismatches 6; Indels 0; Gaps 0;

QY 1 KXSLPELDYERSATEPIYSGQINIEIXYT 28  
DB 3 KXSLPELDYIDYGALEPHINAIQIMQLHHS 30

RESULT 19  
US-09-262-856A-6  
Sequence 6, Application US/09262856A  
Patent No. 633164  
GENERAL INFORMATION:  
APPLICANT: TAKESAKO, Kazutoh  
APPLICANT: MIZUTANI, Shigetoshi  
APPLICANT: ENDO, Masahiro  
APPLICANT: KATO, Ikunoshin  
TITLE OF INVENTION: FUNGAL ANTIGENS AND PROCESS FOR PRODUCING THE SAME  
FILE REFERENCE: 1422-372P  
CURRENT APPLICATION NUMBER: US/09/262,856A  
CURRENT FILING DATE: 1999-03-04  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 6  
LENGTH: 188  
TYPE: PRT  
ORGANISM: Candida albicans  
US-09-262-856A-6

Query Match 49.0%; Score 71; DB 4; Length 188;  
Best Local Similarity 81.2%; Pred. No. 0.00041;

Matches 13; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
Qy 13 ATEPISGOINEIXYT 28  
Db 1 ATEPITCOMNEIHYT 16

RESULT 20  
5240847-24  
Patent No. 5240847  
APPLICANT: HECKL, KONRAD; SPEVAK, WALTER; OSTERMANN, ELINBORG;  
ZOPHEL, ANDREAS; KRISTEK, EDELTRAUD; MAURER-FOGY, INGRID;  
WICHE-CASTANON, MARIA J.; STRATOWA, CHRISTIAN; HAUPTMANN, RUDOLF  
TITLE OF INVENTION: HUMAN MANGANESE SUPEROXIDE DISMUTASE  
(HMN-SOD)  
NUMBER OF SEQUENCES: 34  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/167,261  
FILING DATE: 11-MAR-1988  
SEQ ID NO: 24  
LENGTH: 46  
40847-24

Query Match 48.3%; Score 70; DB 6; Length 46;  
Best Local Similarity 54.5%; Pred. No. 9.6e-05;  
Matches 12; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KYSLPELDYFSPATEPISGOI 22  
Db 25 KHSLPDLAYDYGALPISGKIMELH 46

## RESULT 21

US-08-928-692-30  
Sequence 30, Application US/08928692  
Patent No. 5958727  
GENERAL INFORMATION:  
APPLICANT: Brody, Howard  
APPLICANT: Yaver, Deborah S.  
APPLICANT: Lamsa, Michael  
APPLICANT: Hansen, Kim  
TITLE OF INVENTION: Methods for Modifying the Production of  
TITLE OF INVENTION: a Polypeptide  
NUMBER OF SEQUENCES: 80  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 59587270 No. 59587270disk of No. 5958727th America, Inc.  
STREET: 405 Lexington Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10174  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/928,692  
FILING DATE: 12-SEPT-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Lambiris, Elias J.  
REGISTRATION NUMBER: 33,728  
REFERENCE/DOCKET NUMBER: 4944,200-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-867-0123  
TELEFAX: 212-878-9655  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 230 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

MOLECULE TYPE: No. 5958727e  
US-08-928-692-30

Query Match 47.6%; Score 69; DB 2; Length 230;  
Best Local Similarity 48.1%; Pred. No. 0.0011;  
Matches 13; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

Qy 1 KYSLPELDYFSPATEPISGOINEIXY 27  
Db 36 KATLPDLAYDYGALPISGKIMELH 62

RESULT 22  
US-09-339-972-30  
Sequence 30, Application US/09339972  
Patent No. 6323002  
GENERAL INFORMATION:  
APPLICANT: Brody, Howard  
APPLICANT: Yaver, Deborah S.  
APPLICANT: Lamsa, Michael  
APPLICANT: Hansen, Kim  
TITLE OF INVENTION: Methods for Modifying the Production of  
TITLE OF INVENTION: a Polypeptide  
NUMBER OF SEQUENCES: 80  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 63230020 No. 6323002disk of No. 6323002th America, Inc.  
STREET: 405 Lexington Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10174  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/339,972  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/928,692  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Lambiris, Elias J.  
REGISTRATION NUMBER: 33,728  
REFERENCE/DOCKET NUMBER: 4944,200-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-867-0123  
TELEFAX: 212-878-9655  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 230 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: No. 6323002e  
US-09-339-972-30

Query Match 47.6%; Score 69; DB 4; Length 230;  
Best Local Similarity 48.1%; Pred. No. 0.0011;  
Matches 13; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

Qy 1 KYSLPELDYFSPATEPISGOINEIXY 27  
Db 36 KATLPDLAYDYGALPISGKIMELH 62

RESULT 23  
US-09-091-097-6  
Sequence 6, Application US/09091097  
Patent No. 6432407  
GENERAL INFORMATION:



```
APPLICANT: TAKESAKO, KAZUTOH
APPLICANT: OKADO, TAKASHI
APPLICANT: YAGIHARA, TOMOKO
APPLICANT: KURODA, MASANOBU
APPLICANT: ONISHI, YOSHIMI
APPLICANT: KATO, IKUNOSHIN
APPLICANT: AKIYAMA, KAZUO
APPLICANT: YASUEDA, HIROSHI
APPLICANT: YAMAGUCHI, HIDEYO
TITLE OF INVENTION: ANTIGENIC PROTEIN ORIGINATING IN
TITLE OF INVENTION: MALASSEZIA
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH, LLP
STREET: PO BOX 747
CITY: FALLS CHURCH
STATE: VA
COUNTRY: USA
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/091,097
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: WEINER, MARC S.
REGISTRATION NUMBER: 32,181
REFERENCE/DOCKET NUMBER: 1422-0346P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-205-8000
TELEFAX: 703-205-8050
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 206 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-091-097-6

Query Match          42.8%; Score 62; DB 4; Length 206;
Best Local Similarity 40.7%; Pred. No. 0.014;
Matches 11; Conservative 7; Mismatches 9; Indels 0; Gaps 0;
```

```
1 YSLPELDYFSATEPYISQINEIXY 27
:|||||:|||||:
6 EYLPPLPYAVDALEPFSKEIMTVHH 32
```

RESULT 24

US-09-411-578-1

Sequence 1, Application US/09411578

Patent No. 6203801

GENERAL INFORMATION:

APPLICANT: Schaap, Theodorus C

APPLICANT: Kuiper, Catharina M

APPLICANT: Vermeulen, Arnoldus N

TITLE OF INVENTION: Coccidioides Vaccines

FILE REFERENCE: schaad

CURRENT APPLICATION NUMBER: US/09/411,578

EARLIER FILING DATE: 1999-10-04

EARLIER APPLICATION NUMBER: 98203384.7

EARLIER FILING DATE: 1998-10-07

EARLIER APPLICATION NUMBER: 98203457.1

NUMBER OF SEQ ID NOS: 41

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 1

LENGTH: 214

TYPE: PRT

ORGANISM: Eimeria tenella

US-09-411-578-1

Query Match 40.0%; Score 58; DB 4; Length 214;

Best Local Similarity 46.2%; Pred. No. 0.068;

Matches 12; Conservative 2; Mismatches 12; Indels 0; Gaps 0;

```
QY      2 YSLPELDYFSATEPYISQINEIXY 27
:|||||:|||||:
DB      3 FELPPLPYDALEPFSKETLEHYH 28
```

RESULT 25

US-09-134-001C-3584

Sequence 3584, Application US/09134001C

Patent No. 6380370

GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: GTC-007

CURRENT APPLICATION NUMBER: US/09/134,001C

CURRENT FILING DATE: 1998-08-13/064,964

PRIOR APPLICATION NUMBER: US 60/064,964

PRIOR FILING DATE: 1997-11-08

PRIOR APPLICATION NUMBER: US 60/055,779

PRIOR FILING DATE: 1997-08-14

NUMBER OF SEQ ID NOS: 5674

SEQ ID NO 3584

LENGTH: 201

TYPE: PRT

ORGANISM: Staphylococcus epidermidis

US-09-134-001C-3584

Query Match 39.3%; Score 57; DB 4; Length 201;

Best Local Similarity 42.3%; Pred. No. 0.092;

Matches 11; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

```
QY      2 YSLPELDYFSATEPYISQINEIXY 27
:|||||:|||||:
DB      5 FELPPLPYAVDALEPHDKQMEITHH 30
```

RESULT 26

US-08-679-493A-184

Sequence 184, Application US/08679493A

Patent No. 6303295

GENERAL INFORMATION:

APPLICANT: Taylor, Ethan W.

TITLE OF INVENTION: SELENOPROTEINS, CODING SEQUENCES AND METHODS

FILE REFERENCE: 55-95

CURRENT APPLICATION NUMBER: US/08/679,493A

CURRENT FILING DATE: 1996-07-12

PRIOR APPLICATION NUMBER: 60/001203

PRIOR FILING DATE: 1995-07-14

PRIOR APPLICATION NUMBER: 60/003,112

PRIOR FILING DATE: 1995-09-01

NUMBER OF SEQ ID NOS: 216

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 184

LENGTH: 190

TYPE: PRT

ORGANISM: Entamoeba dispar

US-08-679-493A-184

Query Match 38.6%; Score 56; DB 4; Length 190;

Best Local Similarity 38.5%; Pred. No. 0.12;

Matches 10; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

```
QY      2 YSLPELDYFSATEPYISQINEIXY 27
:|||||:|||||:
DB      3 FOLPPLPYAVNALEPHISKETLEFHH 28
```

## RESULT 27

US-08-445-909A-17  
; Sequence 17, Application US/08445909A  
; Patent No. 5772396  
; GENERAL INFORMATION:  
; APPLICANT: ATKINSON, Antony  
; TITLE OF INVENTION: Pharmaceutical Compositions  
; TITLE OF INVENTION: Containing Superoxide Dismutase from Bacillus  
; TITLE OF INVENTION: Stearothermophilus and Bacillus Caldorena  
; NUMBER OF SEQUENCES: 29  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/445,909A  
; FILING DATE: 22-MAY-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/978,697  
; FILING DATE: 02-FEB-1993  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 204 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-445-909A-17

Query Match 36.6%; Score 53; DB 1; Length 204;  
Best Local Similarity 37.0%; Pred. No. 0.43;  
Matches 10; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

## QY 2 YSLPELDYFSATPEYISQGINEIXYT 28

Db 3 FELPALPYDALBPHDKETMNIHHT 29

RESULT 28  
US-08-445-909A-29  
; Sequence 29, Application US/08445909A  
; Patent No. 5772396  
; GENERAL INFORMATION:  
; APPLICANT: ATKINSON, Antony  
; TITLE OF INVENTION: Pharmaceutical Compositions  
; TITLE OF INVENTION: Containing Superoxide Dismutase from Bacillus  
; TITLE OF INVENTION: Stearothermophilus and Bacillus Caldorena  
; NUMBER OF SEQUENCES: 29  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/445,909A  
; FILING DATE: 22-MAY-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/978,697  
; FILING DATE: 02-FEB-1993  
; INFORMATION FOR SEQ ID NO: 29:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 204 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-445-909A-29

Query Match 36.6%; Score 53; DB 1; Length 204;

Best Local Similarity 37.0%; Pred. No. 0.43;  
Matches 10; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

## QY 2 YSLPELDYFSATPEYISQGINEIXYT 28

Db 3 FELPALPYDALBPHDKETMNIHHT 29

## RESULT 29

US-09-605-785-708  
; Sequence 708, Application US/09605785  
; Patent No. 6321716  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqi  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Darlick  
; APPLICANT: Li, Samuel  
; APPLICANT: Wang, Ajun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Hepler, William  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
; FILE REFERENCE: 210121.427C16  
; CURRENT APPLICATION NUMBER: US/09/605,785  
; CURRENT FILING DATE: 2000-06-27  
; NUMBER OF SEQ ID NOS: 835  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 708  
; LENGTH: 371  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-605-785-708

Query Match 34.5%; Score 50; DB 4; Length 371;  
Best Local Similarity 45.0%; Pred. No. 2.9;  
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

## QY 2 YSLPELDYFSATPEYISQ 21

Db 150 YLPAIDWTSALAPYIGTQ 169

RESULT 30  
US-09-020-956-113  
; Sequence 113, Application US/09020956  
; Patent No. 6261562  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS FOR  
; NUMBER OF SEQUENCES: 178  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED AND BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/020,956
FILING DATE: 09-FEB-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.427C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 113:
SEQUENCE CHARACTERISTICS:
LENGTH: 553 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
-09-020-956-113

```

```

Query Match          34.5%; Score 50; DB 4; Length 553;
Best Local Similarity 45.0%; Pred. No. 4.9;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

```

```

QY 2 YSLPELDYEFSEATPEYISGQ 21
DB 176 YLPAIDWDTSLAPYLGTO 195

```

```

RESULT 31
US-09-030-607-113
Sequence 113, Application US/09030607
Patent No. 6262245
GENERAL INFORMATION:
APPLICANT: Xu, JIANGCHUN
APPLICANT: DILLON, Davin C.
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS FO
NUMBER OF SEQUENCES: 224
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/030,607
FILING DATE: 25-FEB-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.427C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 113:
SEQUENCE CHARACTERISTICS:
LENGTH: 553 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-09-030-607-113

```

```

Query Match          34.5%; Score 50; DB 4; Length 553;
Best Local Similarity 45.0%; Pred. No. 4.9;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

```

```

QY 2 YSLPELDYEFSEATPEYISGQ 21
DB 176 YLPAIDWDTSLAPYLGTO 195

```

```

RESULT 32
US-09-605-785-113
Sequence 113, Application US/09605785
Patent No. 6321716
GENERAL INFORMATION:
APPLICANT: Xu, JIANGCHUN
APPLICANT: DILLON, Davin C.
APPLICANT: MITCHEM, Jennifer L.
APPLICANT: HARLOCKER, Susan L.
APPLICANT: JIANG, Yugu
APPLICANT: HENDERSON, Robert A.
APPLICANT: KALOS, Michael D.
APPLICANT: FANGER, Gary R.
APPLICANT: RETTER, Marc W.
APPLICANT: STOLK, John A.
APPLICANT: DAY, Craig H.
APPLICANT: VEDVICK, Thomas S.
APPLICANT: CARTER, Darrick
APPLICANT: LI, Samuel
APPLICANT: WANG, Aijun
APPLICANT: SKEIKY, Yasir A.W.
APPLICANT: HEPLER, William
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.427C16
CURRENT APPLICATION NUMBER: US/09/605,785
NUMBER OF SEQ ID NOS: 835
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 113
LENGTH: 553
TYPE: PRT
ORGANISM: Homo sapien
US-09-605-785-113

```

```

Query Match          34.5%; Score 50; DB 4; Length 553;
Best Local Similarity 45.0%; Pred. No. 4.9;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

```

```

QY 2 YSLPELDYEFSEATPEYISGQ 21
DB 176 YLPAIDWDTSLAPYLGTO 195

```

```

RESULT 33
US-09-439-313-113
Sequence 113, Application US/09439313
Patent No. 6329505
GENERAL INFORMATION:
APPLICANT: Xu, JIANGCHUN
APPLICANT: DILLON, Davin C.
APPLICANT: MITCHEM, Jennifer L.
APPLICANT: HARLOCKER, Susan Louise
APPLICANT: JIANG, Yugu
APPLICANT: REED, Steven G.
APPLICANT: KALOS, Michael
APPLICANT: FANGER, Gary
APPLICANT: RETTER, Mark
APPLICANT: SOLK, John
APPLICANT: DAY, Craig
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
FILE REFERENCE: 210121.427C9

```

; CURRENT APPLICATION NUMBER: US/09/439,313  
 ; CURRENT FILING DATE: 1999-11-12  
 ; NUMBER OF SEQ ID NOS: 575  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 113  
 ; LENGTH: 553  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapien  
 US-09-439-313-113

Query Match 34.5%; Score 50; DB 4; Length 553;  
 Best Local Similarity 45.0%; Pred. No. 4.9;  
 Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 2 YSLPELDYFSEATEPYISQ 21  
 Db 176 YLLPAIDWTSALAPYLGTQ 195

RESULT 34  
 -09-352-616A-113  
 Sequence 113, Application US/09352616A  
 Patent No. 6395278  
 GENERAL INFORMATION:  
 APPLICANT: Dillon, Davin C.  
 APPLICANT: Harlocker, Susan Louise  
 APPLICANT: Jiang, Yugu  
 APPLICANT: Xu, Jiangchun  
 APPLICANT: Mitcham, Jennifer Lynn  
 TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS  
 FILE REFERENCE: 210121.427C8  
 CURRENT APPLICATION NUMBER: US/09/352,616A  
 CURRENT FILING DATE: 1999-07-13  
 NUMBER OF SEQ ID NOS: 472  
 SOFTWARE: FastSeq for Windows Version 3.0  
 SEQ ID NO 113  
 LENGTH: 553  
 TYPE: PRT  
 ORGANISM: Homo sapien  
 US-09-352-616A-113

Query Match 34.5%; Score 50; DB 4; Length 553;  
 Best Local Similarity 45.0%; Pred. No. 4.9;  
 Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 2 YSLPELDYFSEATEPYISQ 21  
 Db 176 YLLPAIDWTSALAPYLGTQ 195

RESULT 35  
 US-09-602-877A-101  
 Sequence 101, Application US/09602877A  
 Patent No. 6432707  
 GENERAL INFORMATION:  
 APPLICANT: Reed, Steven G.  
 APPLICANT: Xu, Jiangchun  
 APPLICANT: Dillon, Davin C.  
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
 FILE REFERENCE: 210121.446C5  
 CURRENT APPLICATION NUMBER: US/09/602,877A  
 CURRENT FILING DATE: 2000-06-22  
 NUMBER OF SEQ ID NOS: 107  
 SOFTWARE: FastSeq for Windows Version 3.0  
 SEQ ID NO 101  
 LENGTH: 553  
 TYPE: PRT  
 ORGANISM: Homo sapien  
 US-09-602-877A-101

Query Match 34.5%; Score 50; DB 4; Length 553;

Best Local Similarity 45.0%; Pred. No. 4.9;  
 Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 2 YSLPELDYFSEATEPYISQ 21  
 Db 176 YLLPAIDWTSALAPYLGTQ 195

RESULT 36  
 US-09-232-149A-113  
 Sequence 113, Application US/09232149A  
 Patent No. 6465611  
 GENERAL INFORMATION:  
 APPLICANT: Xu, Jiangchun  
 APPLICANT: Dillon, Davin C.  
 APPLICANT: Mitcham, Jennifer Lynn  
 TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE  
 FILE REFERENCE: 210121.427C6  
 CURRENT APPLICATION NUMBER: US/09/232,149A  
 CURRENT FILING DATE: 1999-01-15  
 NUMBER OF SEQ ID NOS: 338  
 SOFTWARE: FastSeq for Windows Version 3.0  
 SEQ ID NO 113  
 LENGTH: 553  
 TYPE: PRT  
 ORGANISM: Homo sapien  
 US-09-232-149A-113

Query Match 34.5%; Score 50; DB 4; Length 553;  
 Best Local Similarity 45.0%; Pred. No. 4.9;  
 Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 2 YSLPELDYFSEATEPYISQ 21  
 Db 176 YLLPAIDWTSALAPYLGTQ 195

RESULT 37  
 US-09-028-934-31  
 Sequence 31, Application US/09028934  
 Patent No. 6117670  
 GENERAL INFORMATION:  
 APPLICANT: Ligon, James M.  
 APPLICANT: Hill, Dwight S.  
 APPLICANT: Lam, Steven T.  
 APPLICANT: Hammer, Philip B.  
 APPLICANT: van Pee, Karl-Heinz  
 APPLICANT: Kirner, Sabine  
 APPLICANT: Young, Thomas R.  
 TITLE OF INVENTION: Pyrolytic Blosynthesis Genes and Uses  
 NUMBER OF SEQUENCES: 37  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: No. 611767artis Corporation  
 STREET: 3054 Cornwallis Road  
 CITY: Research Triangle Park  
 STATE: NC  
 COUNTRY: USA  
 ZIP: 27709  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/028,934  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/729,214  
 FILING DATE: 09-Oct-1996  
 PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/258,261  
 FILING DATE: 08-JUN-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Meigs, J. Timothy  
 REGISTRATION NUMBER: 38,241  
 REFERENCE/DOCKET NUMBER: CGC1506/CIP7  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 919-541-8587  
 TELEFAX: 919-541-8689  
 INFORMATION FOR SEQ ID NO: 31:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 566 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-09-028-934-31

Query Match 33.1%; Score 48; DB 3; Length 566;  
 Best Local Similarity 42.9%; Pred. No. 11;  
 Matches 9; Conservative 6; Mismatches 4; Indels 2; Gaps 1;  
 1 KYSLPELDY--EFSATEPYIS 19  
 71 KYGPELDHITSFYATQRYA 91

RESULT 38  
 US-08-042-747A-8  
 Sequence 8, Application US/08042747A  
 Patent No. 5487969  
 GENERAL INFORMATION:  
 APPLICANT: Eberle, Richard  
 APPLICANT: Black, Darla  
 APPLICANT: Schiniciello, Franco  
 APPLICANT: Hilliard, Julia K.  
 TITLE OF INVENTION: Cloning and Amplification of Monkey B  
 NUMBER OF SEQUENCES: 12  
 CORRESPONDENCE ADDRESSES:  
 ADDRESSEE: Cox & Smith Incorporated  
 STREET: 112 East Pecan Street, Suite 2000  
 CITY: San Antonio  
 STATE: Texas  
 COUNTRY: USA  
 ZIP: 78205  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/042,747A  
 FILING DATE:  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Haymond, W. Bradley  
 REGISTRATION NUMBER: 35186  
 REFERENCE/DOCKET NUMBER: S-0072.179  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 210-554-5500  
 TELEFAX: 210-226-8395  
 TELEX: 767609  
 INFORMATION FOR SEQ ID NO: 8:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 885 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-042-747A-8

QY 2 YSLPELDYFESATEPYISQINE 24  
 DB 582 YSRPLVSFRYADGFLVGGQIGE 604

RESULT 39  
 US-08-804-439A-23  
 Sequence 23, Application US/08804439A  
 Patent No. 6015565  
 GENERAL INFORMATION:  
 APPLICANT: Rose, Timothy M.  
 APPLICANT: Bosch, Marnix L.  
 APPLICANT: Strand, Kurt  
 TITLE OF INVENTION: GLYCOPROTEIN B OF THE RHFV/KSHV  
 TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES  
 NUMBER OF SEQUENCES: 113  
 CORRESPONDENCE ADDRESSES:  
 ADDRESSEE: Fish & Richardson P.C.  
 STREET: 4225 Executive Square, Ste 1400  
 CITY: La Jolla  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 92037  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/804,439A  
 FILING DATE: February 21, 1997  
 CLASSIFICATION: 424  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Haile, Lisa A.  
 REGISTRATION NUMBER: 38,347  
 REFERENCE/DOCKET NUMBER: 09176/004001  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (619) 678-5070  
 TELEFAX: (619) 678-5099  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 23:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 885 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-804-439A-23

Query Match 33.1%; Score 48; DB 3; Length 885;  
 Best Local Similarity 34.8%; Pred. No. 19;  
 Matches 8; Conservative 5; Mismatches 10; Indels 0; Gaps 0;  
 QY 2 YSLPELDYFESATEPYISQINE 24  
 DB 582 YSRPLVSFRYADGFLVGGQIGE 604

RESULT 40  
 US-08-720-229-23  
 Sequence 23, Application US/08720229  
 Patent No. 6023542  
 GENERAL INFORMATION:  
 APPLICANT: Rose, Timothy M.  
 APPLICANT: Bosch, Marnix L.  
 APPLICANT: Strand, Kurt  
 TITLE OF INVENTION: GLYCOPROTEIN B OF THE RHFV/KSHV  
 TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES  
 NUMBER OF SEQUENCES: 100  
 CORRESPONDENCE ADDRESSES:  
 ADDRESSEE: Morrison & Foerster  
 STREET: 755 Page Mill Road

CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304-1018  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/720,229  
FILING DATE: 26-SEP-1996  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Schief, J. Michael  
REGISTRATION NUMBER: 40,253  
REFERENCE/DOCKET NUMBER: 29938-20002.00  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 813-5600  
TELEFAX: (415) 494-0792  
TELEX: 706141  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 885 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-720-229-23

Query Match 33.1%; Score 48; DB 3; Length 885;  
Best Local Similarity 34.8%; Pred. No. 19;  
Matches 8; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 2 YSLPELDYEFKTEPYISQINE 24  
DB 582 YSRPLVSFRYEDGPLVEGQLGE 604

Search completed: April 9, 2003, 14:01:34  
Job time : 7.09174 secs